SYMPOSIUM

TECHNOLOGY, VALUES, AND THE JUSTICE SYSTEM

Technology, Values, and the Justice System: Introduction
Gerry Alexander 1

Washington State Access to Justice Technology Principles 5

Potential Washington State General Court Rule: Access to Justice and Technology 11

Technology and the Washington State Administrative Process—Some Preliminary Notes
William R. Andersen 13

Internet and the Justice System
Vinton G. Cerf 25

Towards a Theory of Legitimate Access: Morally Legitimate Authority and the Right of Citizens to Access the Civil Justice System
Kenneth Einar Himma 31

Donald J Horowitz 77

Conceptualizing the Right of Access to Technology
Morton J. Horwitz 105

Privacy as Contextual Integrity
Helen Nissenbaum 119
Government-to-Citizen Online Dispute Resolution: A Preliminary Inquiry

Anita Ramasastry 159

Rise of the Machines: Justice Information Systems and the Question of Public Access to Court Records over the Internet

Gregory M. Silverman 175

Designing an Accessible, Technology-Driven Justice System: An Exercise in Testing the Access to Justice Technology Bill of Rights

T.W. Small, Robert Boiko & Richard Zorza 223

The Common Law Process: A New Look at an Ancient Value Delivery System

Dennis J. Sweeney 251

Crafting a License To Know from a Privilege To Access

Jane K. Winn 285

Online Court Records: Balancing Judicial Accountability and Privacy in an Age of Electronic Information

Peter A. Winn 307

The End of Technology: A Polemic

Louis E. Wolcher 331

Some Reflections on Long-Term Lessons and Implications of the Access to Justice Technology Bill of Rights Process

Richard Zorza 389

COMMENTS

Keeping Attorneys from Trashing Identities: Malpractice as Backstop Protection for Clients Under the United States Judicial Conference’s Policy on Electronic Court Records

Michael Caughey 407

Balancing Consumer Interests in a Digital Age: A New Approach to Regulating the Unauthorized Practice of Law

Cristina L. Underwood 437

Copyright © 2004, Washington Law Review Association. Except as otherwise provided, the author of each article in this issue has granted permission for copies of that article to be made for classroom use, provided that (1) copies are distributed at or below cost, (2) the author and journal are identified, (3) proper notice of copyright is affixed to each copy, and (4) the Washington Law Review Association is notified of the use.
OFFICERS OF ADMINISTRATION
Lee L. Huntsman, Ph.D., President of the University
W. H. Knight, Jr., B.A., J.D., Dean, Professor of Law
Patricia C. Kaeszler, B.A., M.D., J.D., Associate Dean for Research & Faculty Development, Professor of Law
Penny A. Hazeltin, B.A., J.D., M.L.L., Associate Dean for Library & Computing Services, Professor of Law
Sandra E. Madrid, B.S., Ph.D., Assistant Dean for Student Affairs & Administration
Dexter Bailey, B.S., M.B.A., Assistant Dean for Development
Mary A. Hotchkiss, B.A., M.S.L.S., J.D., L.L.M., Assistant Dean for Academic Services

FACULTY EMERITI
William T. Burke, B.S., J.D., J.S.D.
Harry M. Cross, B.A., J.D.
Robert L. Fletcher, A.B., LL.B.
John Huston, B.A., J.D., L.L.L.
Arval Morris, B.A., M.A., J.D., LL.M., L.L.D.
Cornelius J. Peck, B.S., LL.B.
John R. Price, A.B., LL.B.
Marjorie D. Rombauer, B.A., J.D.
Charles Z. Smith, B.S., J.D.
Frank W. Smith, Jr., B.A., J.D., L.L.M.

FACULTY OF LAW
Craig A. Allen, B.A., J.D., Professor of Law
William R. Andersen, B.S.L., LL.B., LL.M., Judson Falknor Professor of Law
Helen A. Anderson, B.A., J.D., Senior Lecturer
Robert T. Anderson, B.A., J.D., Assistant Professor of Law, Native American Law Center Director
Thomas R. Andrews, B.A., M.A., J.D., Professor of Law
Robert H. Aronson, B.A., J.D., Professor of Law
Kathryn Battuello, B.A., J.D., M.P.H., Research Assistant Professor of Law
Karen Boxx, B.A., J.D., Associate Professor of Law
Steve F. Calandrille, B.A., J.D., Associate Professor of Law
Donald C. Clarke, A.B., Ph.D., J.S.D., Professor of Law
William Covington, B.A., J.D., Assistant Professor of Law
Sam A. Donaldson, B.A., J.D., L.L.M., Assistant Professor of Law
G. Meade Emory, B.A., J.D., L.L.M., Professor of Law, Director of Graduate Program in Taxation
Julia Gold, B.A., J.D., Senior Lecturer
Robert W. Gormlikiewicz, B.A., M.A., J.D., Research Associate Professor of Law, Director of Graduate Program of Intellectual Property Law & Policy
James H. Hardisty, A.B., LL.B., Professor of Law
Gregory A. Hicks, B.A., J.D., Professor of Law
Roland L. Hjorth, A.B., L.L.B., Dean Emeritus, Garvey, Schubert & Barer Professor of Law
Linda S. Hume, B.A., J.D., Professor of Law
Stewart M. Jay, A.B., J.D., William L. Dwyer Professor of Law
John M. Junker, B.A., J.D., Professor of Law
Lisa Kelly, B.A., J.D., Professor of Law
Alan Kirtley, B.A., J.D., Associate Professor of Law, Director of Clinics
Richard O. Kimmert, B.S., M.B.A., LL.B., D. Wayne & Anne Gittinger Professor of Law
Deborah Maranville, B.A., J.D., Professor of Law
Anna Mastroianni, B.S./B.A., J.D., M.P.H., Assistant Professor of Law
Kathleen M. McCann, B.A., J.D., Senior Lecturer
Jacqueline McManus, B.G.S., J.D., Assistant Professor of Law
Peter Nicolas, B.A., M.P.P., J.D., Associate Professor of Law
Sean O'Connor, B.A., M.A., J.D., Associate Professor of Law
Kathleen O'Neill, B.A., J.D., Associate Professor of Law, Director of Legal Research and Writing
Roy L. Prosterman, A.B., LL.B., Professor of Law
Anita Ramasastri, B.A., M.A., J.D., Associate Professor of Law
William H. Rodgers, Jr., B.A., L.L.B., Professor of Law
Eric Schnapper, B.A., M.A., B.Phil., L.L.B., Professor of Law
Scott A. Schumacher, B.A., J.D., Lecturer
William B. Stoebuck, B.A., M.A., J.D., S.J.D., Judson Falkner, Professor of Law Emeritus
Toshiko Takekura, B.A., LL.M., Ph.D., Professor of Law, Director, CASRIP
Veronica L. Taylor, B.A., B.A., L.L.M., Professor of Law, Asian Law Program Director
Michael E. Townsend, B.A., M.A., J.D., Associate Professor of Law
Philip A. Trautman, B.A., J.D., Professor of Law
Lea B. Vaughn, A.B., J.D., Professor of Law
Walter J. Walsh, B.C.L., LL.M. S.J.D., Assistant Professor of Law
Lis W. Wirh, A.B., M.A., J.D., Associate Professor of Law
June K. Winn, B.Sc., J.D., Professor of Law
Ron J. Whitener, B.A., J.D., J.S.D., Professor of Law
Louis E. Wolcher, B.A., J.D., Professor of Law

ADJUNCT AND AFFILIATE FACULTY
G. Andrew H. Benjamin, B.A., M.A., J.D., Ph.D., Affiliate Professor of Law
Daniel M. Bodansky, A.B., Ph.D., Affiliate Professor of Law
Sharon E. Brown, B.A., J.D., Ed.D., Adjunct Research Associate Professor of Law
Daniel H. Foote, A.B., J.D., Affiliate Professor of Law
John O. Haley, A.B., L.L.B., LL.M., Affiliate Professor of Law
Marc J. Hershman, B.A., J.D., Adjunct Professor of Law
Kenneth E. Himma, B.A., M.A., J.D., Ph.D., Adjunct Lecturer of Law
CONTENTS

ARTICLES

Sex and the Workplace: “Consenting” Adolescents and a Conflict of Laws

Jennifer Ann Drobac 471

Wild Dreamers: Meditations on the Admissibility of Dream Talk

Louise Harmon 575

NOTES & COMMENTS

Foster v. Carson: The Ninth Circuit Misapplies the Capable-of-Repetition-yet-Evading-Review Exception to the Mootness Doctrine and Lends a Free Hand to Budget-Cutting State Officials

Joshua C. Gaul 665


Lynette Meachum 693

Capsized by the Constitution: Can Washington State Ferries Meet Federal Screening Requirements and Still Pass State Constitutional Muster?

David J. Perkins 725

Protestors Have Fourth Amendment Rights, Too: In Graves v. City of Coeur d’Alene, the Ninth Circuit Clouds Clearly Established Law Governing Searches

Holly Vance 753
University of Washington School of Law 2003–2004

OFFICERS OF ADMINISTRATION
Lee L. Huntsman, Ph.D., President of the University
W. H. Knight, Jr., B.A., J.D., Dean, Professor of Law
Patricia C. Kauzlar, B.A., M.D., J.D., Associate Dean for Research & Faculty Development, Professor of Law
Penny A. Hazelton, B.A., J.D., M.L.L., Associate Dean for Library & Computing Services, Professor of Law
Sandra E. Madrid, B.S., Ph.D., Assistant Dean for Student Affairs & Administration
Dexter Bailey, B.S., M.B.A., Assistant Dean for Development
Mary A. Hotchkiss, B.A., M.S.L.S., J.D., L.L.M., Assistant Dean for Academic Services

FACULTY EMERITI
William T. Burke, B.S., J.D., J.S.D.
Harry M. Cross, B.A., J.D.
Robert L. Fletcher, A.B., LL.B.
John Huston, B.A., J.D., LL.M.
Arval Morris, B.A., M.A., J.D., L.L.M., J.D.
Cornelius J. Peck, B.S., LL.B.
John R. Price, A.B., LL.B.
Marjorie D. Rombauer, B.A., J.D.
Charles Z. Smith, B.S., J.D.
Frank W. Smith, Jr., B.A., J.D., LL.M.

FACULTY OF LAW
Craig H. Allen, B.A., J.D., Professor of Law
William R. Andersen, B.S.L., LL.B., LL.M., Judson Falknor Professor of Law
Helen A. Anderson, B.A., J.D., Senior Lecturer
Robert T. Anderson, B.A., J.D., Assistant Professor of Law, Native American Law Center Director
Thomas R. Andrews, B.A., M.A., J.D., Professor of Law
Robert H. Aronson, B.A., J.D., Professor of Law
Kathryn Battuello, B.A., J.D., M.P.H., Research Assistant Professor of Law
Karen Boxx, B.A., J.D., Associate Professor of Law
Steve P. Calandrillo, B.A., J.D., Associate Professor of Law
Donald C. Clarke, A.B., M.Sc., J.D., Professor of Law
William Covington, B.A., J.D., Assistant Professor of Law
Sam A. Donaldson, B.A., LL.M., Assistant Professor of Law
G. Meade Emory, B.A., J.D., LL.M., Professor of Law, Director of Graduate Program in Taxation
Julia Gold, B.A., J.D., Senior Lecturer
Robert W. Gormliekiewicz, B.A., M.A., J.D., Research Associate Professor of Law, Director of Graduate Program of Intellectual Property Law & Policy
James H. Hardisty, A.B., LL.B., Professor of Law
Gregory A. Hicks, B.A., J.D., Professor of Law
Roland L. Hjorth, A.B., LL.B., Dean Emeritus, Garvey, Schubert & Barer Professor of Law
Linda S. Hume, B.A., J.D., Professor of Law
Stewart M. Jay, A.B., J.D., William L. Dwyer Professor of Law
John M. Junker, B.A., J.D., Professor of Law
Lisa Kelly, B.A., J.D., Professor of Law
Alan Kirtley, B.A., J.D., Associate Professor of Law, Director of Clinics
Richard O. Kummert, B.S., M.B.A., LL.B., D. Wayne & Anne Gittinger Professor of Law
Deborah Maranville, B.A., J.D., Professor of Law
Anna Mastroianni, B.S./B.A., J.D., M.P.H., Adjunct Research Associate Professor of Law
Kathleen M. McCann, B.A., J.D., Senior Lecturer
Jacqueline McMurtie, B.G.S., J.D., Assistant Professor of Law
Peter Nicolas, B.A., M.P.P., J.D., Associate Professor of Law
Sean O’Connor, B.A., M.A., J.D., Assistant Professor of Law
Kathleen O’Neill, B.A., J.D., Associate Professor of Law, Director of Legal Research and Writing
Roy L. Prosterman, A.B., LL.B., Professor of Law
Anita Ramasasya, B.A., M.A., J.D., Associate Professor of Law
William H. Rodgers, Jr., B.A., LL.B., Professor of Law
Eric Schnapper, B.A., M.A., B.Phil., LL.B., Professor of Law
Scott A. Schumacher, B.A., J.D., Lecturer
William B. Stoebuck, B.A., M.A., J.D., S.J.D., Judson Falknor, Professor of Law Emeritus
Toshiko Takenaka, B.A., LL.M., Ph.D., Professor of Law, Director, CASRIP
Veronica L. Taylor, B.A., B.A., LL.M., Professor of Law, Asian Law Program Director
Michael E. Townsend, A.B., J.D., Ph.D., J.D., Associate Professor of Law
Philip A. Trautman, B.A., J.D., Professor of Law
Lea B. Vaughan, A.B., J.D., Professor of Law
Walter J. Walsh, B.C.L., LL.M., S.J.D., Assistant Professor of Law
Lisa W. Wiehl, A.B., M.A., J.D., Associate Professor of Law
June K. Winn, B.Sc., J.D., Professor of Law
Ron J. Whitner, B.A., J.D., Assistant Professor of Law
Louis E. Wolcher, B.A., J.D., Professor of Law

ADJUNCT AND AFFILIATE FACULTY
G. Andrew H. Benjamin, B.A., M.A., J.D., Ph.D., Affiliate Professor of Law
Daniel M. Bodansky, A.B., M.Phil., J.D., Affiliate Professor of Law
Sharon E. Brown, B.A., J.D., Ed.D., Adjunct Research Associate Professor of Law
Daniel H. Foote, A.B., J.D., Affiliate Professor of Law
John O. Haley, A.B., LL.L., LL.M., Affiliate Professor of Law
Marc J. Hershman, A.B., J.D., Adjunct Professor of Law
Kenneth E. Himma, B.A., M.A., Ph.D., Adjunct Lecturer of Law
Steven G. Olswang, B.A., J.D., Ph.D., Adjunct Professor of Law
Beth E. Rivin, B.A., M.D., M.P.H., Adjunct Clinical Assistant Professor of Law
Hugh D. Spitzer, B.A., J.D., LL.M., Affiliate Professor of Law
Richard O. Zerbe, Jr., A.B., Ph.D., Adjunct Professor of Law
CONTENTS

ARTICLES

Capital Punishment, Proportionality Review, and Claims of Fairness (With Lessons from Washington State)

Timothy V. Kaufman-Osborn 775

Settling Significant Cases

Jeffrey R. Seul 881

NOTES & COMMENTS

First Things First: Federal Courts Should Determine the Legal Status of a Lloyd’s of London Syndicate Before Deciding the Syndicate’s Citizenship for Diversity Purposes

John M. Brust 969

When Prisoners Are Weary and Their Religious Exercise Burdened, RLUIPA Provides Some Rest for Their Souls

Anne Y. Chiu 999

Salient Circumstance Overlooked: The Influence of Parole Search Conditions on Fourth Amendment Analysis in United States v. Kincade

Megan Grembowski 1029
University of Washington School of Law 2003–2004

OFFICERS OF ADMINISTRATION
Mark A. Emmert, B.A., Ph.D., President of the University
W. H. Knight, Jr., B.A., J.D., Dean, Professor of Law
Patricia C. Kuszler, B.A., M.D., J.D., Associate Dean for Research & Faculty Development, Professor of Law
Penny A. Hazelton, B.A., J.D., M.L.L., Associate Dean for Library & Computing Services, Professor of Law
Sandra E. Madrid, B.S., Ph.D., Assistant Dean for Student Affairs & Administration
John McNamara, B.A., Assistant Dean for Development & External Relations
Mary A. Hochkiss, B.A., M.S.L.S., J.D., LL.M., Assistant Dean for Academic Services

FACULTY EMERITI
William T. Burke, B.S., J.D., J.S.D.
Harry M. Cross, B.A., J.D.
Robert L. Fletcher, A.B., LL.B.
John Huston, B.A., J.D., LL.M.
Arval Morris, B.A., M.A., J.D., LL.M., LL.D.
Cornelius J. Peck, B.S., LL.B.
John R. Price, A.B., LL.B.
Marjorie D. Rombauer, B.A., J.D.
Charles Z. Smith, B.S., J.D.
Frank B. Smith, Jr., B.A., J.D., LL.M.

FACULTY OF LAW
Craig H. Allen, B.S., J.D., Professor of Law
William R. Anderson, B.S.L., LL.B., LL.M., Judson Falknor Professor of Law
Helen A. Anderson, B.A., J.D., Senior Lecturer
Robert T. Anderson, B.A., J.D., Assistant Professor of Law, Native American Law Center Director
Thomas R. Andrews, B.A., M.A., J.D., Professor of Law
Robert H. Aronson, B.A., J.D., Professor of Law
Kathryn Battuello, B.A., J.D., M.P.H., Research Assistant Professor of Law
Karen Boxx, B.A., J.D., Associate Professor of Law
Steve P. Calandrillo, B.A., J.D., Associate Professor of Law
Donald C. Clarke, A.B., M.S., J.D., Professor of Law
William Covington, B.A., J.D., Assistant Professor of Law
Sam A. Donaldson, B.A., M.A., J.D., Assistant Professor of Law
G. Meade Emory, B.A., J.D., LL.M., Professor of Law, Director of Graduate Program in Taxation
Julia Gold, B.A., J.D., Senior Lecturer
Robert W. Gomulkiewicz, B.A., M.A., J.D., Research Associate Professor of Law, Director of Graduate Program of Intellectual Property Law & Policy
James H. Hardisty, A.B., LL.B., Professor of Law
Gregory A. Hicks, B.A., J.D., Professor of Law
Roland L. Hjorth, A.B., LL.B., Dean Emeritus, Garvey, Schubert & Barer Professor of Law
Linda S. Hunne, B.A., J.D., Professor of Law
Stewart M. Jay, A.B., J.D., William L. Dwyer Professor of Law
John M. Junker, B.A., J.D., Professor of Law
Lisa Kelly, B.A., J.D., Professor of Law
Alan Kirby, B.A., J.D., Associate Professor of Law, Director of Clinics
Richard O. Kummert, B.S., M.B.A., LL.B., D. Wayne & Anne Gittinger Professor of Law
Deborah Maranville, B.A., J.D., Professor of Law
Anna Mastreanini, B.A., B.A., J.D., M.P.H., Assistant Professor of Law
Kathleen M. McGinnis, B.A., J.D., Senior Lecturer
Jacqueline McMurtrie, B.G.S., J.D., Assistant Professor of Law
Jessica Neilson, B.A., J.D., LL.M., Lecturer, Visiting Scholar Coordinator
Peter Nicolas, B.A., M.P.P., J.D., Associate Professor of Law
Sean O’Connor, B.A., M.A., J.D., Assistant Professor of Law
Kathleen O’Neill, B.A., J.D., Associate Professor of Law, Director of Legal Research and Writing
Roy L. Prosterman, A.B., LL.B., Professor of Law
Anita Ramaswamy, B.A., M.A., J.D., Associate Professor of Law
Michael Robinson-Dorns, B.A., J.D., Associate Professor of Law, Director, Kathy and Steve Berman Environmental Law Clinic
William H. Rodgers, Jr., B.A., LL.B., Professor of Law
Eric Schnapper, B.A., M.A., B.Phil., LL.B., Professor of Law
Scott A. Schumacher, B.A., J.D., Lecturer
William B. Stocker, B.A., M.A., J.D., Professor of Law Emeritus
Toshiko Takeshima, B.A., LL.M., Ph.D., Professor of Law, Director, CASRIP
Veronica J. Taylor, LL.B., B.A., LL.M., Professor of Law, Asian Law Program Director
Michael E. Townsend, B.A., M.A., Ph.D., J.D., Associate Professor of Law
Philipp A. Trautman, B.A., J.D., Professor of Law
Lea B. Vaughn, A.B., J.D., Professor of Law
Walter J. Walsh, B.C.L., LL.M., S.J.D., Assistant Professor of Law
Ron J. Whitener, B.A., J.D., Assistant Professor of Law
Lisa W. Wiehl, A.B., M.A., J.D., Associate Professor of Law
June K. Winn, B.Sc., J.D., Professor of Law
Louis E. Wolcher, B.A., J.D., Professor of Law

ADJUNCT AND AFFILIATE FACULTY
G. Andrew H. Benjamin, B.A., J.D., Ph.D., Affiliate Professor of Law
Daniel M. Bodansky, A.B., M.Phil., J.D., Affiliate Professor of Law
Karen E. Brown, B.A., M.A., J.D., Ed.D., Adjunct Research Associate Professor of Law
Daniel H. Hoete, B.A., J.D., Affiliate Professor of Law
John O. Haley, A.B., LL.B., LL.M., Affiliate Professor of Law
CONTENTS

ARTICLE
Whistling in the Dark? Corporate Fraud, Whistleblowers, and the Implications of the Sarbanes–Oxley Act for Employment Law
Miriam A. Cherry 1029

NOTES & COMMENTS
Does Delaware’s Section 102(b)(7) Protect Reckless Directors from Personal Liability? Only if Delaware Courts Act in Good Faith
Matthew R. Berry 1125

Lucky for Life: A More Realistic and Reasonable Estate Tax Valuation for Nontransferable Lottery Winnings
Kyla C.E. Grogan 1153

Putting Flesh on the Bones of United States v. Winans: Private Party Liability Under Treaties that Reserve Actual Fish for the Tribal Taking
Lindsay Halm 1181

Making Mommies: The Washington State Court of Appeals Exceeded Its Authority by Creating a Common Law Parentage Action in In re Parentage of L.B.
Thomas G. Robinson-O’Neill 1209
2004–2005 EDITORIAL BOARD*

Editor-in-Chief
JOHN BRUST

Managing Editors
KIMBERLY COZZETTO
KYLIA GROGAN
LYNETTE MEACHUM
JILL VALLELY
JUNGMIN JENNIFER YOO

Executive Notes and Comments
Editors
DAVID PERKINS
HOLLY VANCE

Thesis Editors
LINDSAY HALM
TAD ROBINSON–O’NEILL

Notes and Comments Editors
ANNE CHIU
MEGAN GREMBOWSKI
SCOTT HOLLIDAY
KIM MCCLAIN
HEATHER MCKIMME
JENNIFER MURRAY

Associate Editors-in-Chief
DORGEN FRIED
JOSH GAUL

Executive Articles Editors
MATTHEW BERRY
MICHELLE JENSEN

Articles Editors
MICHAEL BEERS
SARAH L. BIRD
EMILY CORDO
KELLY FENNERTY
MELANIE MAYER

Thesis Editors
LINDSAY HALM
TAD ROBINSON–O’NEILL

Notes and Comments Editors
ANNE CHIU
MEGAN GREMBOWSKI
SCOTT HOLLIDAY
KIM MCCLAIN
HEATHER MCKIMME
JENNIFER MURRAY

EDITORIAL STAFF

CHRISTOPHER ABBOTT
JANNA AGINSKY
ROBERT ALLISON
AMANDA CARR
SEAN CROMAN
ERIN CURTIS
JULIE FIELDS

WES HENRICKSEN
TIM HOBBS
EMILY HUTCHINSON
LISA LUEBECK
MELISSA MANKE
DAVID MARTIN
JASON MORGAN

RANDALL OLSEN
JANETE OSLER
MARGARET PAK
ADAM REINS
THOMAS C. SCHROEDER
SARAH SHIREY
MATTHEW STOCK
ERIK VAN HAGEN

*Washington Law Review issues are produced on a calendar-year schedule; the Editorial Board’s term of office, however, corresponds to the school year. Accordingly, editorial responsibilities for November issues are apportioned between consecutive Boards.
Philip A. Trautman, B.A., J.D., Professor of Law
Lea B. Vaughn, A.B., J.D., Professor of Law
Walter J. Walsh, B.C.L., LL.M., S.J.D., Associate Professor of Law
Ron J. Whitener, B.A., J.D., Assistant Professor of Law
Lisa W. Wielh, A.B., M.A., J.D., Associate Professor of Law
Jane K. Winn, B.Sc., J.D., Professor of Law
Louis E. Wolcher, B.A., J.D., Charles I. Stone Professor of Law

ADJUNCT, AFFILIATE, AND VISITING FACULTY
G. Andrew H. Benjamin, B.A., M.A., J.D., Ph.D., Affiliate Professor of Law
Daniel M. Bodansky, A.B., M.Phil., J.D., Affiliate Professor of Law
Sharan E. Brown, B.A., M.A., J.D., Ed.D., Adjunct Research Associate Professor of Law
Michael W. Dowdle, B.M., M.M., M.Phil., J.D., Visiting Associate Professor of Law
Daniel H. Foote, A.B., J.D., Affiliate Professor of Law
John O. Haley, A.B., L.L.B., L.L.M., Affiliate Professor of Law
Marc J. Herschman, A.B., J.D., Adjunct Professor of Law
Nancy S. Jecker, B.A., M.A., Ph.D., Adjunct Associate Professor of Law
David M. Keepnews, B.S., M.P.H., J.D., Ph.D., Affiliate Assistant Professor of Law
Elizabeth F. Loftus, B.A., M.A., Ph.D., Affiliate Professor of Law
Michael McCann, B.A., M.A., Ph.D., Adjunct Professor of Law
Steven G. Olswang, B.A., J.D., Ph.D., Adjunct Professor of Law
Beth E. Rivin, B.A., M.D., M.P.H., Adjunct Clinical Assistant Professor of Law
Hugh D. Spitzer, B.A., J.D., L.L.M., Affiliate Professor of Law
Richard O. Zerbe, Jr., A.B., Ph.D., Adjunct Professor of Law
TECHNOLOGY, VALUES, AND THE JUSTICE SYSTEM: INTRODUCTION*

Gerry Alexander†

On my own behalf and that of my colleagues on the Washington State Supreme Court, I welcome all of you to this conference, which is devoted to the important subject of “Technology, Values, and the Justice System.” We are honored by your presence at this event at which we will together explore developments in information technologies, the use of such technologies in the justice system, and the broader societal ramifications of such use.

In pondering what to say to you in my welcoming remarks, I was struck by the fact that this conference is one of firsts. Let me explain. This is the first major conference to be held at the new home of the University of Washington School of Law, William H. Gates Hall, a building appropriately named for a distinguished lawyer and citizen who was instrumental in the creation of our state’s Access to Justice Board in 1994. I am aware also that this is the first major conference ever dedicated entirely to the subject of technology, values, and the justice system. Finally, this conference will lead to a first ever symposium edition of the *Washington Law Review* devoted to technology and consisting of scholarly writings engendered by this conference.

Another hallmark of this conference is that it is one of collaboration and diversity. On the collaboration side, it is co-hosted by the University of Washington School of Law, the Washington Law Review, and the Access to Justice Technology Bill of Rights Committee of this state’s Access to Justice Board. Co-sponsors are the Information School of the University of Washington and the Shidler Center for Law, Commerce, and Technology. Insofar as diversity is concerned, we have representatives here from all of the aforementioned organizations as well as a large number of other persons who bring a wide variety of perspectives to this event. The program presenters are also diverse. Although we will hear from lawyers and judges, we will also benefit

---

* Remarks made at *Technology, Values, and the Justice System*, a symposium held on January 16–17, 2004 at the University of Washington School of Law.

† Chief Justice, Washington State Supreme Court.
from the views of technologists, ethicists, librarians, philosophers, and historians as well as other individuals who operate at the more practical level of service delivery.

This conference will focus on the fact that whether we like it or not the new information and communication technologies, including the Internet, have begun to enter the justice system, will continue to do so, and will in many ways affect the system in the future. We in the judiciary and other legal fields have come to recognize that the current and future use of such technologies pose significant challenges and opportunities as we continue our quest to guarantee full and equal access to the justice system. Technology can provide increased pathways for access to justice, but it can also perpetuate existing barriers and exclusions and indeed create new barriers. The Washington State justice system is dedicated to ensuring that in the use of such technologies—and elsewhere—barriers to accessing the justice system are avoided, eliminated, or minimized, and that pathways to the justice system are created or maximized.

But if we just let the technology happen and do not act in advance to assure the values of access, inclusion, equality, and quality into the technology and its use, what we desire will not happen, and past deficiencies may persist and worsen. We will not allow that. That is why the Supreme Court’s creation, the Access to Justice Board, fostered the initiative that has brought us here today.

The mission of what has been called the Access to Justice Technology Bill of Rights initiative is to create a body of fundamental principles to ensure that current and future technology both increases opportunities, prevents new barriers, and eliminates or reduces existing barriers to access to and effective use of the justice system, thereby improving the quality of justice for all persons in Washington State. Its further mission is to begin the process of making sure that those principles are made concrete in the daily lives of the people in this state—bringing ideas to reality, principles to practice.

In this process over the last two years, the Access to Justice Technology Bill of Rights initiative has enlisted hundreds of people from all over the state—and many from other parts of the nation—bringing together an amazing diversity of background and thought to inform the process and the products. It was the stimulus for this conference and the symposium. This is the final major public event of the initiative as the products will now evolve from a special project to integrated permanent parts of the justice system, perpetuating the values for which the project and all of us stand.
Introduction

But this conference is not about ending a discussion and drawing firm conclusions as to what to do. It is rather the beginning of further discussions among many, a stimulus to address the essential issues carefully and proactively, and an effort to find over time what we in the justice system seek and, at our best, find—balanced solutions. These are solutions that balance the fundamental values that we possess as Americans and Washingtonians—such as those set forth in our federal and state constitutions. Often that job is by no means easy. Sometimes some of those fundamental values collide with each other, especially in new circumstances where there are new conditions or when, as now, new and potentially powerful tools are available for use. In these new circumstances, we must not simply do business as usual, but must think and imagine and plan and work at how best to find and apply those balanced solutions in the real lives that we are entrusted to serve.

So these are the challenges and opportunities that we face at this conference. We know the program you are about to witness is an outstanding one, and we trust you will benefit from your attendance and will capitalize on what can only be described as an unprecedented opportunity for an exchange of ideas, theories, and opinions.
WASHINGTON STATE ACCESS TO JUSTICE TECHNOLOGY PRINCIPLES

PREAMBLE

The use of technologies in the Washington State justice system must protect and advance the fundamental right of equal access to justice. There is a particular need to avoid creating or increasing barriers to access and to reduce or remove existing barriers for those who are or may be excluded or underserved, including those not represented by counsel.

This statement presumes a broad definition of access to justice, which includes the meaningful opportunity, directly or through other persons: (1) to assert a claim or defense and to create, enforce, modify, or discharge a legal obligation in any forum; (2) to acquire the procedural or other information necessary (a) to assert a claim or defense, or (b) to create, enforce, modify, or discharge an obligation in any forum, or (c) to otherwise improve the likelihood of a just result; (3) to participate in the conduct of proceedings as witness or juror; and (4) to acquire information about the activities of courts or other dispute resolution bodies. Further, access to justice requires a just process, which includes, among other things, timeliness and affordability. A just process also has “transparency,” which means that the system allows the public to see not just the outside but through to the inside of the justice system, its rules and standards, procedures and processes, and its other operational characteristics and patterns so as to evaluate all aspects of its operations, particularly its fairness, effectiveness, and efficiency.

Therefore, these Access to Justice Technology Principles state the governing values and principles which shall guide the use of technology in the Washington State justice system.

Comment to “Preamble”

Access to justice is a fundamental right in Washington State, and the State Supreme Court has recognized and endeavored to protect that right

in its establishment of the Access to Justice Board. From an understanding that technology can affect access to justice, these Access to Justice Technology Principles are intended to provide general statements of broad applicability and a foundation for resolving specific issues as they arise. The various parts of this document should be read as a whole.

A broad definition of the terms used herein is necessary to ensure that our underlying constitutional and common law values are fully protected. The terms used in this document should be understood and interpreted in that light.

These Principles do not mandate new expenditures, create new causes of action, or repeal or modify any rule. Rather, they require that justice system decision makers consider access to justice, take certain steps whenever technology that may affect access to justice is planned or implemented, avoid reducing access, and, whenever possible, use technology to enhance access to justice.

SCOPE

The Access to Justice Technology Principles apply to all courts of law, all clerks of court and court administrators, and to all other persons or parts of the Washington justice system under the rule-making authority of the Court. They should also serve as a guide for all other actors in the Washington justice system.

“Other actors in the Washington justice system” means all governmental and non-governmental bodies engaged in formal dispute resolution or rulemaking and all persons and entities who may represent, assist, or provide information to persons who come before such bodies.

“Technology” includes all electronic means of communication and transmission and all mechanisms and means used for the production, storage, retrieval, aggregation, transmission, communication, dissemination, interpretation, presentation, or application of information.

Comment to “Scope”

This language is intended to make clear that the Access to Justice Technology Principles are mandatory only for those persons or bodies within the scope of the State Supreme Court’s rulemaking authority. It is, however, hoped and urged that these Principles and their values will be applied and used widely throughout the entire justice system.
Access to Justice Technology Principles

It is also intended that the Access to Justice Technology Principles shall continue to apply fully in the event all or any portion of the performance, implementation, or accomplishment of a duty, obligation, responsibility, enterprise, or task is delegated, contracted, assigned, or transferred to another entity or person, public or private, to whom the Principles may not otherwise apply.

The definition of the word “technology” is meant to be inclusive rather than exclusive.

1. REQUIREMENT OF ACCESS TO JUSTICE

Access to a just result requires access to the justice system. Use of technology in the justice system should serve to promote equal access to justice and to promote the opportunity for equal participation in the justice system for all. Introduction of technology or changes in the use of technology must not reduce access or participation and, whenever possible, shall advance such access and participation.

Comment to “Requirement of Access to Justice”

This Principle combines promotion of access to justice through technology with a recognition of the “first, do no harm” precept. The intent is to promote the use of technology to advance access whenever possible, to maintain a focus on the feasible while protecting against derogation of access, and to encourage progress, innovation, and experimentation.

2. TECHNOLOGY AND JUST RESULTS

The overriding objective of the justice system is a just result achieved through a just process by impartial and well-informed decision makers. The justice system shall use and advance technology to achieve that objective and shall reject, minimize, or modify any use that reduces the likelihood of achieving that objective.

Comment to “Technology and Just Results”

The reference to a “just process” reaffirms that a just process is integral to a just result. The reference to “well-informed decision makers” is to emphasize the potential role of technology in gathering, organizing, and presenting information in order that the decision maker
receives the optimal amount and quality of information so that the possibility of a just result is maximized.

3. OPENNESS AND PRIVACY

The justice system has the dual responsibility of being open to the public and protecting personal privacy. Its technology should be designed and used to meet both responsibilities.

Technology use may create or magnify conflict between values of openness and personal privacy. In such circumstances, decision makers must engage in a careful balancing process, considering both values and their underlying purposes, and should maximize beneficial effects while minimizing detrimental effects.

Comment to “Openness and Privacy”

This Principle underlines that the values of openness and privacy are not necessarily in conflict, particularly when technology is designed and used in a way that is crafted to best protect and, whenever possible, enhance each value. However, when a conflict is unavoidable, it is essential to consider the technology’s effects on both privacy and openness. The Principle requires that decision makers engage in a balancing process which carefully considers both values and their underlying rationales and objectives, weighs the technology’s potential effects, and proceed with use when they determine that the beneficial effects outweigh the detrimental effects.

The Principle applies both to the content of the justice system and its operations, as well as the requirements for accountability and transparency. These requirements may mean different things depending on whether technology use involves internal court operations or involves access to and use of the justice system by members of the public.

4. ASSURING A NEUTRAL FORUM

The existence of a neutral, accessible, and transparent forum for dispute resolution is fundamental to the Washington State justice system. Developments in technology may generate alternative dispute resolution systems that do not have these characteristics, but which, nevertheless, attract users who seek the advantages of available technology. Participants and actors in the Washington State justice system shall use all appropriate means to ensure the existence of neutral, accessible, and
Access to Justice Technology Principles

transparent forums which are compatible with new technologies and to
discourage and reduce the demand for the use of forums which do not
meet the basic requirements of neutrality, accessibility, and
transparency.

Comment to “Assuring a Neutral Forum”

Technologically generated alternative dispute resolution (including
online dispute resolution) is a rapidly growing field that raises many
issues for the justice system. This Principle underlines the importance of
applying the basic values and requirements of the justice system and all
the Access to Justice Technology Principles to that area, while clarifying
that there is no change to governing law.

This Principle is not intended in any way to discourage the
accessibility and use of mediation, in which the confidentiality of the
proceeding and statements and discussions may assist the parties in
reaching a settlement; provided that the parties maintain access to a
neutral and transparent forum in the event a settlement is not reached.

5. MAXIMIZING PUBLIC AWARENESS AND USE

Access to justice requires that the public have available
understandable information about the justice system, its resources, and
means of access. The justice system should promote ongoing public
knowledge and understanding of the tools afforded by technology to
access justice by developing and disseminating information and
materials as broadly as possible in forms and by means that can reach the
largest possible number and variety of people.

Comment to “Maximizing Public Awareness and Use”

While assuring public awareness and understanding of relevant access
to justice technologies is an affirmative general duty of all governmental
branches, this Principle expressly recognizes that the primary
responsibility lies with the justice system itself. As stated in the
Comment to the Preamble, none of these Access to Justice Technology
Principles, including this one, mandates new expenditures or creates new
causes of action. At the same time, however, planners and decision
makers must demonstrate sensitivity to the needs, capacities, and where
appropriate, limitations of prospective users of the justice system.
Communicating the tools of access to the public should be done by whatever means is effective. For example, information about kiosks where domestic violence protection forms can be filled out and filed electronically could be described on radio or television public service announcements. Another example might be providing information on handouts or posters at libraries or community centers. Information could also be posted on a website of the Council for Public Legal Education or of a local or statewide legal aid program, using an audible web reader for persons with visual or literacy limitations. The means may be as many and varied as people's imaginations and the characteristics of the broad population to be reached.

6. BEST PRACTICES

To ensure implementation of the Access to Justice Technology Principles, those governed by these principles shall utilize “best practices” procedures or standards. Other actors in the justice system are encouraged to utilize or be guided by such best practices procedures or standards.

The best practices shall guide the use of technology so as to protect and enhance access to justice and promote equality of access and fairness. Best practices shall also provide for an effective, regular means of evaluation of the use of technology in light of all the values and objectives of these Principles.

Comment to “Best Practices”

This Principle is intended to provide guidance to ensure that the broad values and approaches articulated elsewhere in these Access to Justice Technology Principles are implemented to the fullest extent possible in the daily reality of the justice system and the people served by the justice system. The intent is that high quality practical tools and resources be available for consideration, use, evaluation, and improvement of technologies in all parts of the justice system. This Principle and these Access to Justice Technology Principles as a whole are intended to encourage progress, innovation, and experimentation with the objective of increasing meaningful access to quality justice for all. With these goals in mind, the development and adoption of statewide models for best practices is strongly encouraged.
POTENTIAL WASHINGTON STATE GENERAL COURT RULE

ACCESS TO JUSTICE AND TECHNOLOGY*

The Access to Justice Technology Principles appended to this Rule state the governing values, principles, and standards which shall guide the use of technology in the Washington State justice system. These Principles apply to all courts of law, clerks, and court administrators and to all other persons and parts of the Washington justice system under the rule-making authority of this Court. These Principles shall be considered with other governing law and court rules by the courts of the State of Washington in deciding the appropriate use of technology in the administration of the courts and the cases that come before such courts.

Comment:

This Rule does not create or constitute the basis for new causes of action, mandate new expenditures, or repeal or modify any rule. Rather, it requires that justice system decision makers consider access to justice whenever technology is planned, used, or may be used, avoid reducing access to justice, and, whenever possible, use technology to enhance access to justice.

The Access to Justice Technology Principles should also serve as a guide for all other actors in the Washington justice system that are not under the rule-making authority of this Court. These include all governmental and non-governmental bodies engaged in formal dispute resolution or rulemaking and all persons and entities that may represent, assist, or provide information to persons who come before such bodies.

This Rule and the Access to Justice Technology Principles continue to apply fully in the event all or any portion of the performance, implementation, or accomplishment of a duty, obligation, responsibility, enterprise, or task is delegated, contracted, or assigned to another entity.

* This potential Rule was developed by the Access to Justice Technology Bill of Rights Committee and approved by the Access to Justice Board. The Access to Justice Board currently plans to submit it to the Washington State Supreme Court for its consideration.
or person, public or private, to whom this Rule and the Principles may not otherwise apply.

The meaning of the word “use” herein includes but is not limited to planning, design, development, adoption, deployment, and dissemination as well as its ordinary meaning.
TECHNOLOGY AND THE WASHINGTON STATE ADMINISTRATIVE PROCESS—SOME PRELIMINARY NOTES

William R. Andersen∗

I. INTRODUCTION

In today’s world, government regulation and the provision of governmental services require a complex system of administrative agencies. The procedures these agencies follow—cumulatively called “the administrative process”—have been greatly affected by technology as the promise of faster, cheaper, and better ways of managing the process have become apparent. Obvious cost savings can accompany the introduction of technology, such as the massive savings in printing and mailing costs that are possible when material is made available online. Partly, the push for technology has been aided by the perception that the administrative process can be improved in efficiency, comprehensiveness, speed, and transparency. Finally, the expansion of technology in the administrative process is seen by some as enhancing public participation in government through wider and cheaper access to the governmental processes.

This brief report summarizes information obtained by a preliminary survey in the spring of 2003 of what some Washington state agencies are doing with technology. The agencies surveyed included the Utilities and Transportation Commission and the Departments of Ecology, Revenue, Social and Health Services, and Labor and Industries. To keep the subject within bounds—and within the central focus of the Access to Justice Technology Bill of Rights project—the survey inquired into the use of technology in those parts of the administrative process that are similar to legal process generally—i.e., those components of the administrative process through which binding general principles are formulated (rulemaking) and those processes by which individual disputes are resolved (adjudication, licensing, etc.). The survey did not

∗ Judson Falknor Professor of Law, University of Washington School of Law, William H. Gates Hall, Box 353020, Seattle, Washington 98195. E-mail address: ander@u.washington.edu.
consider the many other uses of technology in internal agency management, records control, benefit administration, etc.

A special concern of the survey was the impact of technology on public access to the process. Such impact can be both positive (increasing and broadening the ways members of the public can participate in the process) and negative (limiting the access of some to that process). If significant numbers of an agency’s public are without full access to technology, great care must be taken in adopting new technology. We know from recent studies that while forty percent of low-income Washingtonians have Internet access, the percentage varies around the state from a high in southwestern Washington of fifty-five percent to a low in southeastern Washington of twenty percent. We also know that only about half of those surveyed have home access to the Internet.\(^1\) Obviously, there are many for whom these concerns and issues are real.

Washington state agencies have been among the leaders nationally in the deployment of technology, and much has been done to expand public access to the process. A look at the web sites of the agencies surveyed will demonstrate how far we have come.\(^2\) All of these sites (and those of many other Washington agencies) show that substantial effort has been expended to make access “user friendly.” In Washington, steps have been taken to facilitate access to online material, including guidelines for making web pages accessible to those with physical and visual limitations.\(^3\) Examples of the success of this design include the Department of Revenue’s quick online access to important publications;\(^4\) the Department of Social and Health Services’ Rules and Policies Assistance Unit, which helps interested persons learn about rulemaking

---


and the means to participate in it; the Utilities and Transportation Commission’s primer on agency rulemaking; the Department of Labor and Industries’ Easy Access page that facilitates doing business online; and the Department of Ecology’s capacity to translate its text into several foreign languages.

Work is continuing daily and one hopes it will continue to emphasize simple, comprehensive, and inexpensive access by the public.

II. RULEMAKING

Rulemaking is an obvious place for the extensive use of information technology because of its dependence on widespread exchange of written material. When rulemaking material is put into electronic form, it is capable of almost instantaneous exchange to and from vast numbers of individuals and groups. Included in the written material that may accompany a major rulemaking effort are the internal studies and documents that shaped the proposed rule, the text of the proposed rule along with any explanatory information, rulemaking dockets, calendars and other information about the process, written comments on the draft by possibly thousands of interested persons, and the text of the final rule itself along with a statement of the agency’s supporting reasons and its analysis of public comments on the rule. Where rules are made under procedures that allow hearings and personal appearance of interested parties, there are further opportunities to use information technology in creating and sharing hearing transcripts, and, of course, the hearing process itself may be affected by tele- and video-conferencing technology.

After rules are officially promulgated, agency web sites and e-mail communications can support enforcement and administration of the rules by speeding up access to the text of rules, to interpretive documents and guidelines, and to advisory information to those affected by the rule. Several agencies have taken the advent of the web page as an

opportunity to further efforts at plain language rule writing and user-friendly presentations. One especially good example is the safety and health rules published under the Washington Industrial Safety and Health Act by the Department of Labor and Industries. These user-tested texts and formats show what can be done when accessibility and usability are placed high on the rulemaking priority list. Incidentally, this study did not explore the role of the Code Reviser—the official keeper of rules format standards. One hopes that office is active in (or at least permissive of) these efforts to increase public access to rules.

The survey showed that virtually all of these rulemaking techniques are in use today by Washington state agencies. Some agencies are farther along than others, but all seem to be adapting the rulemaking process to new technology at a rapid rate and responding to targets of opportunity special to the agency and within available agency resources.

There seems to be no centralized direction to this process. The Department of Information Services is in general support of agency efforts but does not have the resources (or has not been given the mission) to be itself an aggressive force for expanding or unifying technology across the system. This bears watching, as will be noted in the final section.

The future will no doubt show more technology-based improvement in agency rulemaking. A number of federal agencies are exploring further rulemaking enhancements, including electronic docket rooms in which all potential commentators—not just those with access to capitol-city paper docket rooms—can learn of the developing discussions about a proposed rule, including what other comments have been filed and sometimes an opportunity to modify one’s own comments in light of other comments filed. Further, there have been calls for fully interactive web-based discussion forums on pending rules, which could greatly enrich the quality of the discussion and even explore areas of agreement.

Technology and the Administrative Process

compromise, and consensus.\textsuperscript{12} Such improvements in discussion could also increase the sensitivity of judicial review of rules as judges can proceed with more certainty about how a rule will actually work in practice. And ex parte contacts, always a problem of perceived fairness in rulemaking, could be reported and opportunity for answering comment made available.\textsuperscript{13}

III. ADJUDICATION

It is harder to generalize about adjudications because of the enormous variety of proceedings that falls within that classification—from the handful of highly complex, formal hearings an agency like the Washington Utilities and Transportation Commission may conduct in a year to the 40,000 relatively informal unemployment cases the Office of Administrative Hearings judges may conduct in that same time span.

All Washington state agencies make heavy use of technology for internal management of the adjudication process, including research and report writing. Especially in complex cases, the time and money savings can be substantial—as those involved in private litigation have discovered.

With respect to dealings with persons outside the agency, the picture is more complex. For the general public, e-mail is widely used to provide information concerning issues involving adjudicated cases. An agency’s web site may include pleadings, testimony, and other non-confidential material generated by adjudications. Additionally, an agency site that makes available the text of decisions in significant prior cases, digests, and indices permits interested members of the public to discern more quickly the agency’s adjudicating policies.

Of course, in on-the-record adjudications, not all contact with members of the public is proper. Forbidden ex parte communications from outside persons to adjudicating officers have been a problem in more than one agency, and systems should be in place to prevent its

\textsuperscript{12} Harvard University and Massachusetts Institute of Technology have established a discussion listserv to link researchers and professionals working on issues related to information technology and the rulemaking process. Harvard-MIT Data Ctr. Hosted Mailing List (2002), \textit{at} http://lists.hmde.harvard.edu/?info=e-rulemaking-l.

\textsuperscript{13} For a review of some rulemaking technology proposals, see generally Barbara H. Brandon & Robert D. Carlitz, \textit{Online Rulemaking and Other Tools for Strengthening Our Civil Infrastructure}, 54 \textit{Admin. L. Rev.} 1421 (2002).
happening or to ensure that should it happen, the improper communication is promptly made public.\textsuperscript{14}

With respect to parties in pending adjudications, the use of e-mail and the web will provide some opportunity to share information, briefs, pleadings, and other documents. However, for a variety of security, technical, and confidentiality reasons, there are limitations on the use of electronic transmissions for formal notices, official filings, and service. Even here, the agencies are working toward the day when most, if not all, of these official actions can be conducted electronically, and this seems an important area to emphasize.

As could have been expected, technology has affected the conduct of the hearing itself, probably more so than in the judicial realm. Beyond the increasingly common laptop computer at the bench and at counsel table, some hearings are supplemented by or conducted entirely through electronic media. Thus, formal hearings at agencies such as the Washington Utilities and Transportation Commission may be supplemented by teleconference, permitting some parties and some witnesses to participate from a distance. Additionally, thousands of the informal unemployment compensation appeals conducted by judges in the Office of Administrative Hearings are conducted entirely by telephone. In hearings where exchanges among the parties and witnesses are important, multi-party teleconference technology is being explored by some agencies.

The use of telephone and video hearings as substitutes for in-person proceedings is becoming as popular as it is controversial. The benefits can be enormous, saving agencies and parties considerable travel time and expense. At the same time, the “virtual” hearings have been criticized as limiting the ability of a hearing officer to control the hearing process (especially when there are numerous parties and witnesses), limiting the ability of parties and witnesses (especially those with little or no English language skills) to understand fully what is happening, and limiting the ability of the factfinder to determine witness veracity.

Incidentally, while video conferences would seem to be better replicas of in-person hearings than are teleconferences, parties and witnesses today seem clearly to prefer the telephone to video as the vehicle for electronic hearings. Agency officials surmise that compared to a video

\textsuperscript{14} \textit{Cf.} WASH. REV. CODE § 34.05.455(5) (2002) (requiring presiding officer who receives ex parte communication to place such communication and responses made, along with other relevant information, on the record of the proceeding).
camera, the telephone is a familiar and non-threatening instrument. The telephone also can be used in the convenience of the witness’s home and, carrying no video information, it does not let personal appearance, dress, or environment become factors in the case. It may be that as video cameras become more common and less threatening, the video hearing will grow in popularity. But from talking to Washington state agencies today, there does not seem to be much enthusiasm for aggressive movement in that direction.

In a telephone hearing, does the physical absence of the witness affect the ability of the factfinder to determine the truth? According to the folklore of the legal profession—embodied in conventional rules of evidence such as the hearsay rule—the traditional answer has been “yes.” Actually seeing the witness has been thought a key element in determining veracity. And most experienced litigators have strong negative reactions to the virtual hearing. Cross examining an inert telephone receiver is just not the same, they assert.

But at least in relatively simple cases, many of Washington’s agency hearing officers seem to be confident that they can judge veracity from telephonic testimony. These judges are, after all, trained factfinders, not common law juries. And they are usually specialists with considerable experience with the kinds of cases they are hearing. Further, it must be noted that the telephone does transmit information about the witness’s tone, inflection, emphasis, and the like—considerably more information about the witness than is revealed by a written account or a third person’s oral account of what a witness said, both of which might be rejected as hearsay in a jury trial. Finally, the telephone witness can be asked follow-up questions for clarification.

Of course, in relatively more complex hearings, in cases where critical testimony is conflicting, or in hearings where interpreters are necessary, electronic substitutes for the in-person hearing can still be troublesome. The Washington Administrative Procedure Act (WAPA) permits electronic hearings in the discretion of the presiding official but requires that no party be prejudiced thereby. The WAPA further requires that each party “have an opportunity to . . . hear, and, if technically and economically feasible, to see the entire proceeding . . . .” The policy reflected in the WAPA—which allows an electronic hearing over the

15. Id. §§ 34.05.410-.494.
16. Id. § 34.05.449(3).
17. Id.
objection of the parties so long as the presiding official thinks it appropriate, and so long as no plausible showing of prejudice can be made—needs to be carefully reviewed as our experience grows. The time may come when allowing the electronic hearing only with the approval of the parties will state the safer policy.

Agency adjudicators can no doubt profit from other work going on under the general heading of online dispute resolution.18

IV. PERMITTING AND LICENSING

Daily, technology seems to be making the processes of getting permits and licenses more efficient. Washington state agencies now permit such things as online filing and payment of taxes, checking of contractor licensing and bonding, renewing contractor licenses, purchasing permits, handling some industrial insurance transactions, renewing car and boat tabs, and even taking a practice driver’s license exam. Joint actions by related agencies are working toward common templates and master permitting systems.

V. COSTS AND BENEFITS OF THE NEW TECHNOLOGY

The benefits have been discussed earlier. They largely involve efficiencies (and lower costs) to the agency and conveniences to those affected in the form of simplicity and speed of access. To some, the benefits are even more than conveniences—not having to appear physically at a hearing may save expense and lost work time and also may be of special value to those with limited means or education.

Reported concerns and costs of implementing new technology (beyond hardware and software costs) involve questions of security and confidentiality, issues of document management, concerns about system failure and, of course, issues related to human adaptation (to screen rather than paper presentation, for example). The agencies are hard at work on security, management, and system reliability issues. How fast progress can be made on the human factors side is harder to predict. After all, both agency personnel and their clients or customers have long (and largely satisfactory) experience with paper technology that has not

---

Technology and the Administrative Process

changed much since the days of Gutenberg, and this ballast may impede the rate of growth somewhat.

A special cost is borne by those without the equipment or the skills necessary to use the new technology. Government (especially the administrative process) deals with many “ordinary” people who may lack the ability, training, or equipment to permit full access to a system that is highly technologically oriented. Importantly, the number of agency clients and customers who do have access seems to be growing, though not in all cases by dramatic increments. Most of the agencies surveyed expressed concern about this problem. Some have affirmative public education programs such as the Environmental Hearings Office’s online handbook for pro se parties.¹⁹ Others judge that the vast majority of their clients or customers have access to the Internet, either by direct access or through public facilities such as library computers. The use of computers in public libraries, of course, presumes the affected persons have a library in some proximity and that they possess the skills necessary to communicate in this fashion. It seems that access is most likely when the party is a large business, less likely when the party is a very small business or individual, and especially unlikely when the party is an individual with limited means, education, or other limitations.

For those without access, agencies continue to make available (or to receive) paper documents, which can be delivered by mail or obtained through visits to agency offices. Washington state agencies report that they are generous with free copies of paper materials. What remains is the question about the degree to which paper filing puts a claimant, a commentator, or an applicant at some disadvantage in time, in completeness, or even in appearance. More generally, we now and then see reports of states abandoning paper distributions to the public in favor of electronic transmissions, usually as a cost saving move.²⁰ These kinds of developments need to be carefully watched.

---


VI. CONCLUDING OBSERVATIONS

Not everyone is a fan of pushing technology further into the administrative process. In rulemaking, especially, which has become so burdensome and slow as to earn the adjective “ossified,” adding still further process steps and further opportunities for broad public involvement needs careful evaluation. Large federal agencies sometimes receive hundreds of thousands of comments on a proposed rule, and overwhelming agencies with voluminous and duplicative public commentary would not seem in anyone’s interest. Still, it is difficult to argue against wide participation if participation itself is thought valuable in drafting sound rules. We really need better techniques for dealing with redundancy in commentary.

A theoretical concern is the claim that at some point, widespread public commentary on pending rules blurs the important distinction between direct and representative democracy. Of course, a lot depends on the discipline of agency officials (and on judges permitting officials to exercise that discipline). Wisdom cannot depend alone on a show of hands, and most rulemaking statutes permit agencies to make judgments that do not depend solely on quantitative expressions of the public’s preference. If agencies continue to seek the public interest in qualitative terms—and if judges permit them to do this—technology should not unduly impact our representative democracy.

Note that on the question of accessibility, the problem has two aspects: impediments resulting from conventional physical handicaps, such as limited vision, hearing, dexterity, etc.; and impediments arising from lack of equipment or skills. Both are serious problems. On the first, consultants in the states are beginning to design more easily accessible sites. In Washington, steps have been taken to facilitate access to online material, including guidelines for web page design. On the second, the question is: Is enough being done to ease access for those who are not handicapped in the conventional sense but who have limited ability to use technology because of inadequate equipment or skill?

24. See STYLEGUIDE, supra note 3.
Finally, the absence of a central agency for promoting and designing technology for Washington state agencies has plusses and minuses. On the plus side, lack of central control leaves agencies free to experiment. It also allows them to adjust the rate of their own technological progress to their own needs and resources and to the preferences of their constituents. On the minus side, lack of a more centralized effort means unevenness in agency exploration of technology. Further—and this is very important for some members of the public—lack of central design control will surely mean a different look and feel for technology interfaces across the agencies; this may steepen the on-ramp to the technology highway for those unfamiliar with technology generally.

It is hoped that directly or as a model for adaptation and use, the state’s administrative agencies will consider the Access to Justice Technology Bill of Rights as one set of principles and techniques for dealing with problems of the kind addressed here.
INTERNET AND THE JUSTICE SYSTEM

Vinton G. Cerf

Preface: When my friend, Don Horowitz, invited me to comment on the relationship of the Internet to the justice system, I became alarmed, not having seriously considered that this amazing instrument of information propagation might have negative as well as positive effects on the achievement of justice. Judge Horowitz will have done our society a great service simply by asking the question and insisting that we address it.

I. INTRODUCTION

In a fundamental way, our system of justice is rooted in the sharing of information. Every court decision contributes to a growing body of interpretations of the law and influences future decisions in concrete ways. One is reminded of the Jewish Torah and its associated commentaries. These commentaries, stretching over hundreds of years, help to inform and illuminate the interpretation of the writings and guidance found in the Torah. As our society evolves towards increasing connectivity and online being, our reliance on access to accumulated information and wisdom increases. The Internet and its World Wide Web are early harbingers of the information-dense future to come.

In his well-received book, Technologies of Freedom, Ithiel de Sola Pool highlighted the sociopolitical effects of certain, now familiar technologies such as copiers, fax machines, telephones, and mass media. That the Internet is one such technology can hardly be in doubt. The network provides for a freedom of expression that goes far beyond conventional mass media. In the latter, only a select few have the ability to speak with an amplified voice. Perhaps more important, the Internet and the World Wide Web provide for the opportunity to hear as well as

---

* Vinton Cerf serves as Senior Vice President for Technology Strategy for MCI and as Chairman of the Board of the Internet Corporation for Assigned Names and Numbers. He is the co-creator of the TCP/IP protocols and the Internet architecture.

1. The Talmud including the Mishnah, and some wags think also of the Meshuginnah.


to speak. It is the access to such a diverse range of information sources that gives the Internet its unique power.

But the Internet and the World Wide Web are also indiscriminate amplifiers of virtually any information. Whether the information is valid or not, the Internet faithfully transports it to web browsers, chat rooms, instant messaging clients, and streaming video receivers. It is this egalitarian character that makes the Internet a conundrum in practice. On the surface, one cannot draw any foregone conclusions about the information found on the Internet. One can use digital signatures to preserve and, if desired, identify the source of information on the Internet. But the accuracy is only as good as the binding of identity to an individual or source.

In this brief Essay, it is my intention to outline some of the implications of widespread access to and reliance on the Internet with respect to our American system of justice.

II. THE TYRANNY OF INFRASTRUCTURE

When some system, service, or facility truly becomes a part of our infrastructure, we simply use it without thinking much about it until it stops working. Anyone who has experienced the inconvenience of an extended power outage or a long period of congestion on a normally uncongested highway will appreciate this thought. Infrastructure is simply something that is there to be relied upon implicitly. Of course, it doesn’t really work that way. One has to work very hard to make sure that parts of our civil infrastructure actually function reliably.

The Internet is becoming such an infrastructure. While it is not exactly everywhere, it has penetrated our American society beyond 40–50% of American households and businesses to say nothing of university and college campuses, and primary and secondary schools. It has even become a factor in American political life. It is estimated that there are about 187 million users of the Internet in North America and about 650–700 million users worldwide.4

To the extent that any significant part of our system of justice relies on the operation of the Internet, any infirmities in the network, or the applications that rely on it, will be amplified and reflected by impairing that part of the justice system that relies on it. If we deliberately build

---

into the system the presumption of reliable operation, then it is incumbent on us to assure that we are not building on weak foundations. We must take pains to take into account any dependencies and to assure that there are alternatives if the Internet service is unavailable.

Accessibility is another aspect of dependence, and it too must be addressed. If the Internet plays a key role in the archiving and propagation of information in our system of justice, then it must be accessible to all who are seeking justice. Access means more than simply the ability to use a computer that is connected to the Internet—for example, at a public library. Access means the information is organized and presented in such a way that someone with a motor, vision, hearing, or other impairment that would inhibit convenient access to the Internet is accommodated. This could mean tools for voicing Web content and email, or captioning for audio segments that would be inaudible to a deaf person, or tools for simplifying mouse and keyboard interactions for persons with motor impairments. We have some considerable distance to go before the content of the Internet is uniformly accessible, but we can at the least insist that content critical to the justice system is as accessible as we can make it.

III. PRECISION AND RECALL

In the world of library science, there are two important search concepts: precision and recall. What one wants is information that is precisely what we are looking for and the assurance that we have found all the information that is precisely relevant to our search. Failure in either dimension produces incomplete or irrelevant information that is inimical to utility.

In a system of justice that depends heavily on precedent, searching for relevant prior decisions is utterly fundamental. To the extent that the Internet is becoming, or could become, a primary repository of information concerning court decisions at all levels, it is critically important to have search and indexing tools that produce highly precise and complete responses. Otherwise, we run a serious risk of basing our legal arguments on incomplete or weakly relevant positions. That we have such a risk today with the existing framework of indices and tools in the offline world does not excuse an ambition to offer better tools and results in the online world.
Indeed, using the tools of today’s Internet, a competent attorney must not make the assumption that a casual Google search is sufficient to obtain all relevant information. As good as the search engines may be today, they still fall far short of ideal. The information present on the Internet is still relatively unstructured and rendered less easily searched than one would like. Technology is on the way in the form of Extended Markup Language (XML) and the notion of the Semantic Web that the author of the original World Wide Web, Tim Berners-Lee, is now pursuing. By properly annotating information stored on the Web, searching can be made to produce far more relevant and precise retrieval than has been possible with generally unstructured information.

Nor can one make any deep assumptions about the accuracy of information found on the Internet. One must take into account the sources of information, and even that can be a challenge because it is so easy to put misinformation onto the Internet or to mislabel its origins. The implication is that we need to take authentication far more seriously than we currently do. The use of digital signatures and registration of cryptographic certificates can go a long way towards documenting the provenance of information on the Internet, but the infrastructure for such registration and subsequent use to validate the source and integrity of the information is still very immature.

Digital signature technology was first invented around 1977 by Martin Hellman and Whitfield Diffie, both then at Stanford University. They speculated on the existence of mathematical functions that would support the concept of two complementary cryptographic keys—one for encrypting and the other one for decrypting. That this notion sounds counterintuitive compared to systems in which the same key is used for both encrypting and decrypting is precisely why their contribution is so powerful. In essence, it is possible to use this two-key system to “sign” a digital object in such a way that everyone can verify the signature using a public key, but no one but the holder of the secret key can produce the signature.

Once one is provided with the public key for validating digitally signed information, one might wish to look up the key in a public directory to ascertain the source of the signature (and the guarantor of

Internet and the Justice System

the validity of the information). Of course, one has to then be assured that the binding of the public key and its holder is accurate and has not itself been deliberately falsified!

IV. CRITICAL THINKING

No amount of technology will replace the value and importance of critical thinking. Any information obtained on the Internet through the World Wide Web, email, or many other information applications must be subject to critical thinking as to its accuracy, source, and validity. One must be aware of the many ways in which information might become corrupted or polluted in such an open environment as the Internet.

It was once the case that people thought that information printed in a book was likely to be true—why else would someone have gone to the trouble to print it? We know, of course, that people may well go to the trouble of printing books, newspapers, or magazines that carry misinformation—either by intent or simply by accident. It is so easy to put information onto the World Wide Web or to produce electronic mail for distribution on the Internet that one must be intensely aware of the potential for misinformation to be circulated in the system. Anyone who has received a strident email announcing some kind of virus that requires users to remove a particular piece of software from their personal computers will appreciate the power of hoaxes in the online world.

To make matters even more complicated, information that is placed on the Web with the best of intentions, and the highest quality and accuracy, can be polluted by hackers who break into the computers hosting the information and add, change, or delete critical parts of the information content. This extends to software as well and increases the risk of downloading a piece of “Trojan horse” software that may do what it advertises, but may also perform functions (such as sending all your passwords to a web site) that you do not wish to have done.

All of these hazards simply underscore the importance of thinking critically about any information received through the Internet or through any other means and taking pains to validate its source and accuracy.

V. EFFECTS OF BAD INFORMATION

Bad information propagates as fast or faster than good information. Anyone who has had an incorrect credit report or who has experienced identity theft can appreciate how quickly and deeply a misinformation infection can spread. Confusion as to the identity of a convicted felon,
whose name is not unique, can have devastating and lasting negative effects. Once information pollution has occurred, correcting the problem is not unlike trying to put toothpaste back into the tube through the small dispensing end! Incorrectly documented court decisions could have ripple effects throughout the justice system and these potential hazards must be guarded against and repaired rapidly if discovered.

It is sometimes said the antidote for bad information is not censorship but more information. The Internet offers a unique opportunity in the annals of communications for the rapid correction of misinformation. For the first time, it is possible to respond to bad information in the same medium in which it propagated, thereby allowing the correcting information to be found using the same searching tools that might discover the bad information.

Of course, some bad information may be put into the system with the deliberate intent to do harm. This is all the more difficult to deal with if the source of the misinformation is unknown. Anonymity is a valuable commodity under some circumstances, but it can be a serious barrier to the correction of bad information if the source is disguised. Plainly, there are circumstances in which anonymity is in fact quite important (e.g., whistle-blower laws), but it is fair to say that trust in the validity of information can be significantly enhanced if its source can be accurately identified.

VI. CONCLUSION

It should be readily apparent that to the extent that the Internet plays a key role in the archiving and propagation of information associated with our justice system, it is vital that this information be as accurate and complete as possible, that its origins be verifiable, and that the system in which this information is kept is as secure and reliably accessible as possible. The powerful tools at our disposal for the organization and propagation of information have the potential to improve the quality of our justice system and the practice of those who are responsible for the dispensation of justice. But, as this brief Essay suggests, these benefits will not come without significant effort to assure the integrity of the information provided and to assure that the system housing and delivering information is reliably available at need.
Towards a Theory of Legitimate Access: Morally Legitimate Authority and the Right of Citizens to Access the Civil Justice System

Kenneth Einar Himma*

It is undeniable that the new information technologies of the last few decades have been put to many public uses that increase citizen access to the civil justice system. Nearly every state, for example, has published its statutory codes on official websites that can easily be found through legal portals, making the content of such codes readily accessible to citizens from the convenience of their homes. Similarly, many state agencies publish online legal forms that were formerly available only at certain physical locations, such as a courthouse or municipal building. Incorporating these technologies into the legal system has made it much easier for many citizens to access the civil justice system.

Nevertheless, such technologies, if improperly used, can instead diminish the access of many citizens to the civil justice system. One obvious concern is that many citizens lack effective access to the new information technologies because they either cannot afford such access or lack the necessary education to take advantage of them. Devoting limited public resources to publishing public documents exclusively on the web, for example, can effectively exclude such persons from access to those documents. Another concern is that some persons with access to the new technologies may not be able to take full advantage of them. Information published using streaming video technology, for example, is not available to many disabled persons who rely on text readers unable to read video.

Given that these technologies can be used in ways that potentially harm citizens by decreasing their access to the civil justice system, it is reasonable to think that there are general moral principles that govern the

---

* I am extremely indebted to Adrienne Cobb and R. Lee Sims for their generous help with some of the research and to Marsha Iverson, Public Relations Specialist for the King County Library System.

state’s use of such technologies. At the most abstract level, these principles constrain state behavior by defining a general right on the part of citizens to access the civil justice system. The permissibility of incorporating any given technology into the civil justice system would be at least partly determined by whether it violates the general right of access that is defined by these principles.

This Article considers the issue of what the state is morally obligated to provide by way of citizen access to the civil justice system. It begins by describing the general problem of morally legitimate authority and how it bears on the problem of access to the civil justice system. It then identifies three different approaches to the general problem of morally legitimate authority and argues that none of these approaches warrants thinking that the state is morally obligated to provide each citizen with perfectly equal access to the civil justice system. The argument concludes that the three approaches to legitimacy converge on two principles: one that defines an affirmative obligation (the Reasonable Access Principle) to provide to each citizen what is minimally necessary to develop and defend a plausible legal position, and one that defines a negative obligation (the Equality Principle) to refrain from restricting access to the civil justice system for reasons that deny the equality of every moral person.

Three observations about the character and scope of the argument and thesis would be helpful. First, the thesis is not that the two principles identified exhaust the scope of the state’s obligations regarding citizen access to the civil justice system; it is rather that these two principles define necessary conditions for moral legitimacy in this regard. Second, as the title suggests, the argument purports to be no more than a step “towards” a theory of legitimate access and is hence somewhat schematic in character; a fully adequate defense of the Reasonable

---

2. Presumably, any act that can potentially harm innocent persons is subject to the requirements of morality—though these requirements may differ according to whether the agent is an individual or a state. See infra Part I for more discussion on this point.

3. Since there may be other relevant moral principles, the claim that a particular use does not violate the general right of access does not imply that it is morally permissible; the use may violate other principles.

4. This approach has an advantage for skeptics of reductive approaches to theorizing about legitimacy. If one believes, as I do, that there is a plurality of basic principles that jointly define the conditions of legitimacy (rather than just one principle), then an analysis showing that three plausible candidates for inclusion in a non-reductive theory of legitimacy converge on two principles goes a long way towards making the case that any plausible non-reductive theory of legitimacy will include these two principles.
Towards a Theory of Legitimate Access

Access and Equality Principles would require more space than is available here. Third, the Article does not discuss the new information technologies except where helpful to illustrate a particular issue. While the various ways in which the new technologies can be misused highlight the need for a general theory of legitimate access, the principles defining the foundation for such a theory can be developed without discussing these misuses. Indeed, as a logical matter, the uses to which a state might put these technologies present application-contexts for the two principles of legitimate access; such principles apply to these examples and are not justified by these examples.

I. THE PROBLEMS OF ACCESS AND MORALLY LEGITIMATE AUTHORITY

A. The Centrality of Access to the Justice System in the Protection of Legal Rights

Legal systems regulate behavior through the promulgation of legal rules and principles. These legal rules and principles typically purport to establish duties that require subjects to behave or refrain from behaving in certain ways. Sometimes these duties are unconditional (though not necessarily absolute), as in the case of the criminal rule that prohibits assault and in the case of the tort rule that requires us to protect other persons from reasonably foreseeable injuries that might proximately result from our behavior. Sometimes these duties are conditional, as in the case of the general contract rule that requires us to honor the terms of agreements into which we freely enter.

In many instances, these legal duties purport to give rise to legal rights. X's legal duty to refrain from assaulting Y is fairly characterized as defining a legal right on the part of Y against X that X not assault Y.  

5. The issue of whether legal systems have some unique defining function is controversial among legal philosophers. Lon L. Fuller believed that a definition of law must include the idea that the law’s essential function is to “achiev[e] . . . [social] order . . . through subjecting people’s conduct to the guidance of general rules by which they may themselves orient their behavior.” Lon L. Fuller, A Reply to Professors Cohen and Dworkin, 10 VILL. L. REV. 655, 657 (1965). Other philosophers deny that law has any unique conceptual function, arguing that legal systems can be deployed to achieve a variety of purposes. See, e.g., Stephen Perry, The Varieties of Legal Positivism, 9 CAN. J.L. & JURISPRUDENCE 361 (1994).

6. Wesley Newcomb Hohfeld argued that, as a conceptual matter, claim-rights correlate with duties or obligations in the following way: “if X has a right against Y that he shall stay off the former’s land, the correlative (and equivalent) is that Y is under a duty toward X to stay off the
Likewise, $X$’s legal duty to take reasonable measures to protect $Y$ from reasonably foreseeable injuries that might proximately result from $X$’s behavior is fairly characterized as defining a legal right on the part of $Y$ against $X$ that $X$ not negligently injure $Y$. Indeed, the recognition, creation, or establishment of legal rights, if not conceptually essential to the existence of a legal system, is utterly central to legal practice in legal systems that resemble ours in theoretically salient respects.

However, the recognition, creation, or establishment of a legal right by statutory or judicial promulgation has little prudential value to the putative holders if the legal system does not provide some mechanism for addressing violations of those rights. To protect a legal right in a prudentially meaningful way, the law must do two things. First, and most obviously, it must attach some legal consequence to the violation of a right: the law must stipulate either that the violating party is subject to some requirement that self-interested rational agents are likely to regard as undesirable or that the aggrieved party is eligible for some form of relief that self-interested rational agents are likely to regard as desirable. Second, it must provide some mechanism by which the person whose right is violated may activate those consequences against the person who violates that right.

Though legal theorists tend to focus on the first of these features, it is clear that both are needed to endow a legal right with even minimal prudential value. In worlds like ours, where the interests of rational self-interested agents frequently conflict, a right that can always be violated without any legal consequences whatsoever has no prudential value whatsoever to the holder of the right. But this is no less true of a right that fails to provide some mechanism by which the right-holder can
activate those consequences. No matter how undesirable the legal consequences of violating a right might be to the offending party, those consequences cannot engage the offending party’s prudential interests unless there is some mechanism for activating them. Similarly, no matter how desirable those legal consequences might be to the aggrieved party, those consequences cannot engage the aggrieved party’s prudential interests unless there is some mechanism for activating them.

The problem of access, then, is utterly central to the practices of any legal system that purports to establish rights with prudential value to rational self-interested agents—no less so than the problem of assigning consequences to violations. Accordingly, the question of what sort of mechanisms a legal system ought (as a matter of utility, fairness, or justice) to provide for aggrieved parties to activate these consequences is no less central to normative legal theory and political morality than the question of what sorts of consequences ought (as a matter of utility, fairness, or justice) to be attached to the violation of various rights.

B. The Problem of Morally Legitimate Authority

The problem of justifying state authority arises because the state enacts laws that purport to dictate the behavior of those who reside within the physical boundary of the state. As Joseph Raz aptly puts the matter:

A government does not merely say to its subjects: “Here are our laws. Give them some weight in your considerations. But of course you may well be justified in deciding that on balance they should be disobeyed.” It says: “We are better able to decide how you should act. Our decision is in these laws. You are bound by them and should follow them whether or not you agree with them.”

There are two features of moral concern here. The first is that the state purports to preempt the citizens’ own judgments about what they ought to do by issuing directives the citizens are required to obey regardless of whether they want to do so. The second is that the legal directives of the state are backed with coercive measures that are impermissible in any other context: it is hard to think of any other context in which one person or set of persons can legitimately incarcerate or execute another

10. See, e.g., JOSEPH RAZ, ETHICS IN THE PUBLIC DOMAIN (1994) (especially ch. 10).
person. The need to justify state authority arises, then, because the state issues preemptive directives that are enforced with measures that are commonly regarded as presumptively wrong.

While the problem of state authority is especially acute in the criminal context where the violation of a legal rule is frequently linked to lengthy periods of incarceration, the state’s authority in civil contexts also stands in need of moral justification. If the availability of coercive mechanisms is most conspicuous in the criminal context, such mechanisms also play a central role in civil litigation: other things being equal, the point of bringing a civil lawsuit for a plaintiff is to secure a coercive order from the court that requires the defendant to do something. In tort law, the plaintiff seeks a coercive order of money damages. In contract law, the plaintiff seeks a coercive order for either money damages or specific performance. It is fair to say, however, that any plaintiff who brings a civil suit in any legal system that even remotely resembles this one is asking the court not only for a judgment, but also—and equally importantly—for a court order enforcing that judgment.

Indeed, behind every judicial order lies the power to impose sanctions for contempt of court. Contempt sanctions can be civil or criminal in nature. What distinguishes civil and criminal contempt is not the conduct giving rise to the sanction; both cases involve conduct that, in some way, obstructs the administration of justice pursuant to law. What distinguishes civil and criminal contempt is the purpose for which the sanction is imposed. Civil contempt is imposed to induce a party to comply with a particular order of the court and is hence primarily remedial in nature. Criminal contempt is imposed to punish a party for behavior, and not to induce compliance with court orders; though a punitive measure may have the secondary effect of inducing a recalcitrant subject to comply with the order, its primary purpose remains punitive.

In either case, though, the tools available to the court are the same. Civil and criminal contempt sanctions may include incarceration as well as fines, and may include even the power to coerce enforcement.

Towards a Theory of Legitimate Access

from other relevant persons, such as the attorneys for the relevant parties.\textsuperscript{17} It is this power that enables judges to enforce their orders even in civil cases where they cannot plausibly be characterized as imposing direct or indirect sanctions. Without the contempt sanction, judges would have, at the very most, indirect means for enforcing orders requiring payment of civil damages, specific performance, or civil injunctions.

Insofar as state coercion is thus central to enforcing the criminal and civil law, its role must be justified in both of those contexts. The state’s authority to incarcerate a defendant in a civil suit for failing to obey a court order is no less problematic from the standpoint of political morality than its authority to incarcerate a criminal defendant for violating a criminal statute. The application of coercive force to people in any context requires some sort of theoretical account that shows it is justified; to the extent that people have a presumptive moral claim to autonomously direct their own behavior, the state’s exercise of coercive authority over the individual comes into direct conflict with such a moral claim and hence requires moral justification.

\textbf{C. Theories of Legitimacy and Access to Justice}

There are a variety of theories that attempt to specify the conditions under which the state may legitimately enforce its laws. Some of these theories are both comprehensive and reductive in the sense that they attempt to justify all the coercive practices of the state by reference to one or two general moral principles. Contractarian theories, for example, argue that the state’s general authority to coercively dictate behavior is morally justified because citizens either have consented to such authority\textsuperscript{18} or would consent to it under ideal conditions.\textsuperscript{19} Since, on this line of analysis, autonomous moral agents can create binding obligations for themselves by means of consensual agreements (ideal or actual), such agreements are morally sufficient to justify the state’s use of coercive mechanisms, in effect, to enforce laws that are posited pursuant to the terms of the agreement.

Other theories focus on particular areas of law, attempting to show that the use of coercive mechanisms to enforce those areas can be morally justified. For example, John Stuart Mill’s influential claim that

\textsuperscript{17} See id. § 7.21.030(2)(c)–(d).
\textsuperscript{18} See, e.g., JOHN LOCKE, SECOND TREATISE OF GOVERNMENT ch. 8, § 95 (Thomas P. Peardon ed., Bobbs-Merrill Co. 1952) (1690).
\textsuperscript{19} See, e.g., JOHN RAWLS, A THEORY OF JUSTICE 10 (2d ed. 1999).
the state can justifiably prohibit only those acts that are “harmful to others” specifies the limits of the state’s legitimate lawmaking authority to criminalize behavior.\textsuperscript{20} Normative theories of punishment supplement such accounts by showing that the state’s authority to criminalize behavior is legitimately backed by the power to incarcerate and possibly execute those persons who engage in proscribed behavior.\textsuperscript{21}

Theories that justify the state’s authority in areas of civil law tend to be more specific, focusing on the content of a particular substantive area of law, rather than on a general capacity of the state to enact civil regulations of behavior. For example, Charles Fried justifies the content of the contract law by showing it coheres, for the most part, to uncontroversial moral principles governing promises.\textsuperscript{22} Jules Coleman justifies the content of tort law by showing that it embodies a legitimate conception of corrective justice.\textsuperscript{23} Robert Nozick argues that property law is justified to the extent that it coheres with moral principles that define a natural right to property, which includes the authority to freely alienate one’s interest in property by a variety of morally effective consensual mechanisms.\textsuperscript{24}

While there are a number of theories attempting to describe the conditions of procedural legitimacy, they have largely focused on the criminal context. For example, these theories focus on the procedural rights that a criminal defendant should have, which include (but are not limited to) a right to a fair trial, a right to competent representation, a right to appeal, and a right to be acquitted if the evidence does not meet the “beyond a reasonable doubt” standard.\textsuperscript{25} Normative theories of criminal procedure attempt to describe the principles that determine whether a trial is “fair,” representation is “competent,” and so on.

The general issue of what kind of access to the civil justice system a state must provide in order to be legitimate has received comparatively little attention from normative legal and political theorists. While legal


\textsuperscript{21} For a helpful summary of the various normative theories of punishment, see Toni M. Massaro, Shame, Culture, and American Criminal Law, 89 Mich. L. Rev. 1880 (1991).


\textsuperscript{24} ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA 150–53 (1974).

Towards a Theory of Legitimate Access

theorists have devoted considerable space to critically evaluating the requirements of various rules and principles of civil procedure, the focus of such efforts is different from the general normative issue of how much access to civil justice is minimally needed for state legitimacy. In the typical case, the issue is whether a particular rule or interpretation of the rule is consistent with the constitutional requirements of due process—which, for all practical purposes, simply assumes that the United States Constitution defines the morally appropriate standards of access. For this reason, one can criticize the rules of civil procedure as being too lenient or too restrictive without developing a general theoretical account of how much access is minimally consistent with the legitimacy of a legal system. The assumption that the Constitution defines legitimate standards of access to civil justice provides an implicit standard for evaluating the rules of civil procedure that ostensibly circumvents the need for a general theory of legitimate access.

If academic lawyers have paid little attention to such issues, legal and political philosophers have paid even less. A search of the Philosopher’s Index, which is the most comprehensive database of abstracts for philosophical books and articles, for the phrase “access to justice” failed to turn up even one abstract; in contrast, the phrase “distributive justice” turned up 435 abstracts. While searches for the terms “civil liability” and “due process” yielded thirty abstracts, none had anything to do with the issue of general access to the justice system in civil cases. In addition, not one of sixteen leading anthologies for courses in philosophy of law and political philosophy featured even one article on

26. It is important to note that such issues include many that are implicated by the more general issue. For example, a number of articles have been written on the subject of whether courts should waive filing fees and other financial impediments in cases of indigent plaintiffs. See, e.g., John MacArthur Maguire, Poverty and Civil Litigation, 36 HARV. L. REV. 361 (1923); Harry P. Stumpf, Law and Poverty: A Political Perspective, 3 WIS. L. REV. 694 (1968). Such discussions typically assume, without argument, more general principles of legitimacy that are applied to the specific facts. In contrast, my concern here is to argue for these general principles of legitimacy by grounding them in three general approaches to theorizing about moral legitimacy.


29. APPLIED SOCIAL AND POLITICAL PHILOSOPHY (Elizabeth Smith & H. Gene Blocker eds., 1994); JURISPRUDENCE CLASSICAL AND CONTEMPORARY: FROM NATURAL LAW TO POSTMODERNISM (Robert L. Hayman, Jr., et al. eds., 2d ed. 2002); JURISPRUDENCE: TEXT AND READINGS ON THE PHILOSOPHY OF LAW (George Christie & Patrick Martin eds., 2d ed. 1995); JUSTICE AND ECONOMIC DISTRIBUTION (John Arthur & William Shaw eds., 2d ed. 1991); LAW AND MORALITY: READINGS IN LEGAL PHILOSOPHY (David Dyzenhaus & Arthur Ripstein eds., 1996); LLOYD’S INTRODUCTION TO JURISPRUDENCE (M.D.A. Freeman ed., 7th ed. 2001); THE NATURE
the subject. Although such observations are anecdotal in character, they seem pretty clearly to show that the general problem of access has gotten comparatively little attention from philosophers.  

Given the centrality of access to the legitimate operation of institutions that purport to establish and recognize rights, this is a truly remarkable omission. Since, as I argued above, the establishment and recognition of legal rights can have no prudential value whatsoever to citizens in the absence of a mechanism that enables them to activate certain kinds of legal consequences when those rights have been violated, adequate access to the justice system is essential to ensure that legal rights have prudential value to citizen rights-holders. If it is true, as classically liberal theories of legitimacy typically assume, that the normative point of state authority is to serve the interests of citizens by establishing and recognizing rights, then adequate access to the civil justice system is a necessary condition for moral legitimacy. The problem of access is no less central to theories of legitimacy than the

\[\text{AND PROCESS OF LAW (Patricia Smith ed., 1993); PHILOSOPHY OF LAW (Joel Feinberg & Jules Coleman eds., 7th ed. 2004); PHILOSOPHY OF LAW (Conrad Johnson ed., 1993); THE PHILOSOPHY OF LAW (Frederick Schauer & Walter Sinnott-Armstrong eds., 1996); PHILOSOPHY OF LAW AND LEGAL THEORY (Dennis Patterson ed., 2003); POLITICAL PHILOSOPHY; ESSENTIAL SELECTIONS (Aeon J. Skoble & Tibor R. Machan eds., 1999); READING IN THE PHILOSOPHY OF LAW (John Arthur & William H. Shaw eds., 3d ed. 2001); READING IN THE PHILOSOPHY OF LAW (Keith Culver ed., 1999); SOCIAL AND POLITICAL PHILOSOPHY (George Sher & Baruch A. Brody eds., 1999); SOCIAL IDEALS AND POLICIES: READINGS IN SOCIAL AND POLITICAL PHILOSOPHY (Steven Luper ed., 1999).}\]

30. Even the most eminent political philosophers seem to overlook the problem of legitimate access. For example, John Rawls’s principles of justice are intended to guide the “basic structure of society,” which he describes as follows:

\[\text{[T]he basic structure of society is the way in which the main political and social institutions of society fit together into one system of social cooperation, and the way they assign basic rights and duties and regulate the division of advantages that arises from social cooperation over time. The political constitution with an independent judiciary, the legally recognized forms of property, and the structure of the economy (for example, as a system of competitive markets with private property in the means of production), as well as the family in some form, all belong to the basic structure. The basic structure is the background social framework within which the activities of associations and individuals take place. A just basic structure secures what we may call background justice.}\]

JOHN RAWLS, JUSTICE AS FAIRNESS: A RESTATEMENT 10 (2001). The omission of any mention of the importance of access to the justice system to the basic structure of society by arguably the most eminent political philosopher in the last one hundred years speaks volumes about the lack of attention philosophers have paid to the issue.

31. See supra Part I.A.

32. For the leading discussion of this conception of authority, see, for example, RAZ, supra note 10, ch. 10. The remainder of this Article will assume that the moral point of a state is to serve its subjects.
Towards a Theory of Legitimate Access

problem of identifying the substantive constraints on legitimate lawmaking activities.

For this reason, a comprehensive understanding of moral legitimacy requires an understanding of the conditions of morally legitimate access. Since no state can be fully legitimate without guaranteeing adequate access to the civil justice system, it follows that a fully comprehensive theory of legitimacy must specify the conditions that are minimally necessary for morally legitimate access to the civil justice system. No general theory of legitimacy, then, can succeed without including a theory of morally legitimate access. The absence of a general theory of morally legitimate access leaves a profound gap in our understanding of legitimate state authority.33

Moreover, the foregoing considerations suggest an important adequacy constraint on general theories of legitimate access. If the moral point of the state is to serve the prudential interests of its subjects by establishing and protecting rights, then the state must ensure that its practice with respect to rights is sufficient to endow those rights with prudential value to the rights-holders. Since a legal right cannot have any prudential value unless it provides a mechanism that enables rights-holders to activate consequences that conduce to their prudential interests, it follows that a state must provide minimal access to such mechanisms to assure that its legal rights are prudentially valuable. This requirement, then, defines an important adequacy constraint on general theories of morally legitimate access: no theory of legitimate access can be adequate from the standpoint of classical liberalism without assuring sufficient access to the civil justice system to guarantee the prudential value of legal rights to all rationally self-interested citizens.

II. THE STRONG PRINCIPLE OF EQUAL ACCESS

It is tempting to think that the question of how much access to the justice system a legal system is obligated to provide in civil matters, as a matter of political morality, has a straightforward answer: perfectly equal access. On this view, the principle that all citizens have utterly equal moral worth, a principle presumably assumed by every plausible theory

33. A gap that subsequent parts of this Article attempt to close.
of legitimacy, implies that every citizen should have utterly equal access to
the justice system in civil disputes.\(^{34}\)

While it is difficult to state the principle without vagueness, what this Article calls the Strong Principle of Equal Access (SPEA) can be expressed somewhat more rigorously as follows. According to SPEA, the legal system should take affirmative measures to ensure that every two citizens have \textit{exactly} the same level of access to whatever official resources are available for the judicial redress of grievances in the civil context. If one citizen \(X\) has greater access to some resource or mechanism \(m\) of the civil justice system than another \(Y\), then an injustice has occurred that the state must remedy by providing whatever is needed to ensure that \(Y\) has the same access to \(m\).\(^{35}\) Otherwise put, the state must take affirmative steps to ensure that no one person’s access to the civil justice system is more difficult or burdensome than any other person’s access.

This Part of the Article argues that satisfaction of SPEA is not a necessary condition for moral legitimacy. To this end, it considers three different theories of legitimacy that might be adduced in support of SPEA and argues that none provides a foundation for this principle. While perfectly equal access might ultimately be ideal in a utopian world where material resources are unlimited, such access is not required in this world where the quantity of resources that can be deployed by the state in any one area is limited by material conditions of scarcity.

\textbf{A. The General Principle of Equality}

The principle that affirms our moral equality as human beings is grounded in our status as moral persons.\(^{36}\) Each of us has certain properties, capacities, and potentialities that confer the inviolable moral status of personhood. This status comes with a full-blown array of

\(^{34}\) The idea that citizens have utterly equal worth should not be construed as inconsistent with the claim that some moral agents have morally good character and deserve praise for their actions and some have morally bad character and deserve blame, censure, or punishment.

\(^{35}\) Strictly speaking, the inequality can be remedied by diminishing the access of \(X\), rather than increasing the access of \(Y\). But, from the standpoint of morally legitimate access, this guts SPEA of any assurance of even minimal access to the justice system. On this purely formal construction of the principle, a state of affairs in which no one has any access to any mechanisms that would facilitate the resolution of civil disputes would be consistent with SPEA. It is clear that, if it is to define meaningful constraints on what a legitimate state may do at all, SPEA must assume that someone has minimal access to such mechanisms. What follows assumes such a construction.

\(^{36}\) Again, this should not be construed in a manner inconsistent with our ordinary judgments of character. See \textit{supra} note 34.
Towards a Theory of Legitimate Access

general moral rights typically thought to include rights to life, liberty, and property; indeed, as Mary Anne Warren defines the notion, to be a moral person is simply to be a full-fledged member of the moral community.\textsuperscript{37} To the extent that each person has a full set of moral rights, it follows that every human being is entitled to equal respect. The general principle of equality is ultimately a consequence of more theoretical claims governing humanity and personhood.

There is little dispute that states, as well as individuals, are constrained by the moral principle of equality. As a matter of political morality, states have no more liberty than any citizen to violate (as opposed to infringe) a person’s right to life.\textsuperscript{38} Of course, it may be true that states can permissibly kill persons in circumstances where individuals cannot; if, for example, capital punishment is morally legitimate, it seems clear that only the state can permissibly execute someone. But, assuming the permissibility of capital punishment, this is not because a person’s right to life does not give rise to obligations on the part of the state; the state is no less obligated to respect the life of an innocent person than is any individual. It is rather that a person may be executed, as a general moral matter, only under circumstances that can be ensured by the state (or something approximating a state); a suitably reliable finding of guilt in circumstances that afford defendants a fair chance to defend themselves is, presumably, a necessary prerequisite to the legitimate administration of the death penalty.

It is tempting to think that SPEA can be straightforwardly deduced from the principle of equality. If every person has a full and equal set of rights that bind both the state and individuals, then the state must provide every person with a full and equal right to access the justice system in the civil context. But this implies that the state is morally obligated to ensure that every two citizens have exactly the same access to the civil justice system. On this line of reasoning, then, SPEA is simply a corollary of the principle of equality.


\textsuperscript{38} By definition, to say that a right has been “infringed” is to say only that someone has acted in a way that is inconsistent with the holder’s interest in that right; strictly speaking, then, the claim that a right has been infringed is a purely descriptive claim that connotes no moral judgment as to whether or not the infringement is wrong. In contrast, to say that a right has been “violated” is to say that the right has been infringed by some act and that the relevant act is morally wrong. Accordingly, it is a conceptual truth that it can be permissible for an individual or entity to infringe a right, but not to violate a right.
As it turns out, however, this very natural inference is problematic. The claim that every person has a full set of equal moral rights does not imply that the state must, as a moral matter, ensure perfect equality in every morally relevant respect. Every known state, for example, allows considerable inequality in the distribution of material resources among its citizens. While some political theorists believe that such inequalities are justified to the extent that they conduce to overall utility, others believe that such inequalities are justified only to the extent that they conduce to everyone’s advantage. But every classically liberal political theorist believes that a substantial number of these inequalities can, as a matter of political morality, legitimately be permitted by the state.

Here it is crucial to note that inequalities in wealth are not inert with respect to the exercise of the very rights to which personhood gives rise. Persons with more wealth will, for example, be better able to exercise their expressive rights in politically sensitive contexts than persons with less wealth. Indeed, it is not only easier for wealthy persons in democratic societies to win a position in government; it is easier for them to run in the first place. Running for an elected office may require the payment of a substantial filing fee and other expenses that make it comparatively more difficult for less affluent persons. While it is reasonable to think that there are moral limits on how much inequality in this regard may permissibly be allowed by the state, it is also reasonable to think that some such inequality is permissible if it is permissible for the state to allow theoretically significant inequality in the distribution of resources.

The same is true for access to the civil justice system. Strictly speaking, a strong principle of equal access, like SPEA, requires the state to take affirmative steps to ensure that it is no more difficult for one person to access the justice system in a civil context than it is for any other person. But if, as a logical matter, the state can allow substantial inequality in the distribution of resources needed to exercise the right to free speech consistently with the principle of equality, it follows that, as a logical matter, the state can allow some inequality of access to the civil justice system consistently with the principle of equality. If, for example, it is legitimate to require filing fees as a prerequisite for filing a civil

40. See, e.g., RAWLS, supra note 19.
41. Indeed, as a matter of definition, a theory is “classically liberal” only insofar as it purports to legitimate the basic structure of a constitutional democracy that protects private property and other basic rights.
Towards a Theory of Legitimate Access

lawsuit, then it will be somewhat more difficult for less affluent persons who must pay such fees\textsuperscript{42} to access the civil justice system than for more affluent persons.

To the extent, then, that the principle of equality is grounded in the status of every human being as a moral person, it does not logically imply that the state must take affirmative steps to ensure perfectly equal access to either elected office or to the civil justice system. The status of personhood confers rights to life, liberty, and property that are equal to those of any other person. But it does not confer a right to conditions that ensure that it is no more difficult to exercise those rights than it is for any other person. Insofar as the state is obligated to provide equal protection of the rights associated with moral personhood, this obligation, by itself, does not entail an obligation to provide utterly equal access to the civil justice system. Though it is clear that there are moral limits on the extent to which the state may permit inequality of access to the civil justice system, it is equally clear that some inequality of access is logically consistent with the general principle of equality as grounded in principles regarding moral personhood. If this is correct, then SPEA is not a logical consequence of the general principle of equality.

B. Utilitarian Theories of Legitimacy

According to utilitarianism, the moral value of any act is fully determined by its effect on net aggregate utility among members of the community.\textsuperscript{43} Utilitarian moral theories posit a particular state of affairs as objectively good (i.e., the promotion of aggregate utility) and define an act as morally right to the extent that it promotes this favored state of affairs (i.e., to the extent that it promotes aggregate utility) and morally wrong to the extent that it fails to promote this favored state of affairs. Since an act’s effect on utility is an extrinsic feature of the act,\textsuperscript{44} utilitarian theories presuppose that the moral quality of an act does not depend on its intrinsic (or inherent) features and hence that no act is

\textsuperscript{42} The qualification “who must pay such fees” accounts for the possible case where the state is morally obligated to waive filing fees in the case of severely impoverished citizens.

\textsuperscript{43} Utility is usually defined in subjective terms of happiness, pleasure, or well-being. As we will see in this Part, this definition creates epistemic difficulties in evaluating acts under utilitarianism.

\textsuperscript{44} An act can have radically different consequences depending on the circumstances of its performance. For example, whether the act of giving a medication to someone promotes utility depends on to whom the medication is given. Whereas giving chemotherapy, which is highly toxic, to cancer patients can improve their utility, giving it to healthy persons can worsen their utility by increasing the probability that they develop certain kinds of cancer in the long-term.
inherently good or inherently bad. Acts are good (or bad) only insofar as they conduce (or fail to conduce) to the utility of members in the community.\textsuperscript{45}

As a general moral theory, utilitarianism applies both to acts of individuals and to acts of the state. Applied to the state, it implies that the state’s lawmaking authority is constrained by a duty to enact laws that maximally promote aggregate community utility. As political theorist Henry Sidgwick phrases the point:

\begin{quote}
[T]he true standard and criterion by which right legislation is to be distinguished from wrong is conduciveness to the general “good” or “welfare.” And probably the great majority of persons would agree to interpret the “good” or “welfare” of the community to mean, in the last analysis, the happiness of the individual human beings who compose the community; provided that we take into account not only the human beings who are actually living but those who are to live hereafter . . . .

Accordingly, . . . the happiness of the persons affected [is] the ultimate end and standard of right and wrong in determining the functions and constitution of government.\textsuperscript{46}
\end{quote}

Utilitarian theories of legitimacy, then, assess acts of the state entirely in terms of whether they sufficiently conduce to the favored state of affairs (i.e., maximal promotion of utility among the citizenry). The state’s sole obligation, on this view, is to act in ways that have the effect of maximally promoting net utility among its citizens.\textsuperscript{47}

Like any other proposed state measure, the legitimacy of SPEA is thus fully determined, according to utilitarian theories of legitimacy, by whether it adequately promotes the utility of its citizens. Accordingly, the test for assessing SPEA under utilitarian theories of legitimacy is whether, other things being equal, a state that takes affirmative measures

\textsuperscript{45} This distinguishes consequentialist theories like utilitarianism from deontological theories, which assert that the moral quality of some acts is determined entirely by their intrinsic features. On this view, for example, lying is intrinsically wrong—and hence wrong regardless of whether it happens to promote community utility.

\textsuperscript{46} Henry Sidgwick, \textit{Utility and Government}, in \textit{SOCIAL AND POLITICAL PHILOSOPHY}, supra note 29, at 35.

\textsuperscript{47} John Stuart Mill argued that considerations of utility justified the general principle that the state can legitimately prohibit only those acts that are harmful to others. On Mill’s view, utility is most likely to be maximized in a society where people are free to develop and act on their own conceptions of the good; people who are allowed to pursue their own values and plans are more likely to develop the sorts of skills and abilities that will make them useful to other people. Mill, supra note 20, chs. II–III. See generally \textit{JOHN STUART MILL, UTILITARIANISM} (Oxford U. Press 1998) (1871) (providing a general account of Mill’s utilitarianism).
Towards a Theory of Legitimate Access

to ensure that no one person has greater ability than any other to access the civil justice system is more likely to maximize utility among its citizens than a state that permits some inequality of access. If not, then it would be illegitimate for the state to adopt SPEA as a principle of access.

Unfortunately, the task of evaluating SPEA is complicated by the fact that there are serious epistemic difficulties involved in assessing the legitimacy of complex institutional practices under utilitarianism. One difficulty is that our ability to determine in advance what the consequences of any complex practice will be is highly limited because there are a variety of social conditions that can causally interact with elements of the practice in unpredictable ways; the more possible variables there are, the more difficult it is to reliably determine which of a substantial number of outcomes is the most likely.48 A second difficulty is that it is just not clear how to go about making the interpersonal utility comparisons that would have to be made in order to properly evaluate a complex institutional practice. To determine whether I should treat John to his favorite meal or Jane to her favorite meal, I would have to determine whether John gets more enjoyment out of his favorite meal than Jane gets out of hers, which requires a direct comparison of John’s and Jane’s subjective mental states—something to which only John and Jane have direct access. As Robert Goodin explains the problem, “[i]nsofar as utility refers essentially to a state of mind . . . , taking a utility reading requires me to get inside someone else’s head.”49 Because of such difficulties, a utilitarian analysis of SPEA will have to be limited to a comparatively rough assessment of the most salient possible consequences; for this reason, the analysis in this subpart will be somewhat more schematic and speculative than the analysis in the preceding Part.50

48. For a general discussion of such difficulties, see HEIDI HURD, MORAL COMBAT (1999) (especially ch. 8).
50. In this connection, it is worth noting that even the most famous of utilitarian arguments are typically schematic in these respects. For example, Mill justifies his view that the only legitimate criminal laws are those that prevent harm to others (i.e., the so-called Harm Principle) largely on the strength of an identification of three possible consequences of allowing freedom to act in ways that do not harm others: such freedom (1) conduces to the development of a person’s rational faculties; (2) is psychologically satisfying; and (3) conduces to debate that increases the likelihood of discovering truth. While these are, of course, obvious possibilities, it is notable that Mill’s discussion does not (and could not, given the epistemic difficulties) go much beyond identifying three of a fairly large number of possible outcomes. See MILL, supra note 20, chs. II–III.
As a first step towards evaluating SPEA, it is helpful to note that legal norms permitting attorneys to represent clients on a contingency basis provide the poor with substantial access to the civil justice system. The availability of representation on a contingency basis helps to ensure that less affluent plaintiffs can obtain adequate legal representation in cases involving serious injury that allegedly results from a defendant’s breach of duty. The ability of attorneys to recover contingency fees that, strictly speaking, exceed the costs of their services calculated according to a reasonable hourly rate enables them to accept cases that they could not otherwise accept because of their uncertain outcomes. The permissibility of contingency fees thus allows less affluent plaintiffs with economically significant injuries to access the civil justice system in cases where the expected outcomes are comparatively uncertain. Accordingly, the rules permitting contingency fees increase net aggregate community utility by promoting the utility to less affluent plaintiffs.

It is reasonable to think that net aggregate community utility could be further increased by transferring some material resources from the most affluent citizens to the least affluent citizens to improve the latter’s access to the civil justice system. Recent studies suggest that the poorest citizens in the United States are underserved by the civil justice system. For example, a 1994 survey by the American Bar Association disclosed, among other things, that more than fifteen percent of low income

51. For example, The Washington Rules of Professional Conduct provide that “[a] fee may be contingent on the outcome of the matter for which the service is rendered.” WASH. RULES OF PROF’L CONDUCT R. 1.5(c) (2003). The rules further provide that

[a] contingent fee consisting of a percentage of the monetary amount recovered for a claimant, in which all or part of the recovery is to be paid in the future, shall be paid only (i) by applying the percentage to the amounts recovered as they are received by the client or (ii) by applying the percentage to the actual cost of the settlement or award to the defendant.

Id. 1.5(c)(2).

52. Indeed, Philip Corboy argues that contingency fees are morally justified in virtue of increasing the access of less affluent persons to the justice system. Philip H. Corboy, Contingency Fees: The Individual’s Key to the Courthouse Door, LITIG., Summer 1976, at 27; see also Peter Karsten, Enabling the Poor To Have Their Day in Court: The Sanctioning of Contingency Fee Contracts, a History to 1940, 47 DEPAUL L. REV. 231 (1998).

53. As a general matter, whether or not an attorney can justify accepting a contingency case on economic grounds is determined, in part, by the probability of favorable outcomes and the percentage of the recovered amount that the attorney can permissibly charge. The higher the percentage, the stronger the economic justification for taking the case. See infra Part II.C (detailing the calculation of “expected values” in prudential decision-making); see also Thomas J. Miceli & Kathleen Segerson, Contingent Fees for Lawyers: the Impact on Litigation and Accident Prevention, 20 J. LEGAL STUD. 381 (1991) (providing an economic analysis of contingency fees).
Towards a Theory of Legitimate Access

households have a civil legal problem that they cannot afford to formally pursue;\(^{54}\) if the results of the survey are correct, then a significant number of the 34.5 million Americans who live in households with incomes below the poverty level are still not being adequately served by the civil justice system.\(^{55}\) Given that the poorest citizens must struggle to satisfy their basic needs while the richest can satisfy their most extravagant wants, one can plausibly argue that the disutility associated with the material difficulties experienced by the poorest citizens in attempting to enforce their rights outweighs the utility to the wealthiest of the material resources that would collectively be needed to alleviate some of these difficulties.

But it is also reasonable to think that the public costs associated with alleviating each of the inconveniences experienced by all but the wealthiest citizens outweigh the public benefits associated with providing such citizens with easier access to the civil justice system. Insofar as SPEA requires eliminating every relative inconvenience in accessing the civil justice system, it will require compensating less affluent citizens for a host of costs not incurred by the richest citizens. SPEA would require, for example, that the state compensate the poor for whatever wages they lose in pursuing a lawsuit because wealthier citizens commonly receive their income in salaries that are guaranteed and can hence take time off from work, when needed, to pursue a lawsuit without being financially penalized by their employers. Similarly, SPEA would require that the state compensate the poor for any miscellaneous expenses they incur in accessing the justice system, such as unusual transportation costs or the costs of paying for adequate childcare. Though no studies provide a reliable basis for predicting the direct costs of such measures, it is reasonable to hypothesize that these costs are not trivial.

Taking these steps, however, is likely to result in further costs to taxpayers by increasing the number of lawsuits and thereby adding to a growing burden on the civil justice system. Because the costs associated with pursuing a lawsuit are not insubstantial, they likely deter a significant number of lawsuits among less affluent would-be plaintiffs.\(^{56}\)

---

56. This, of course, is true of any material good: the more it costs, the less likely any particular buyer is likely to purchase it.
Reducing these costs and their deterrent effects would make it considerably more likely that would-be plaintiffs—at all income levels—litigate their disputes. This would result in a significant increase in judicial workload that would likely require more facilities and more personnel—resources that taxpayers would ultimately have to subsidize.

Some of these increases are undoubtedly justified from a utilitarian standpoint. If, for example, the costs associated with pursuing a lawsuit deter many less affluent individuals from initiating suits that they would win against more affluent defendants, reducing those costs would increase overall utility in the community. Since the marginal utility of money diminishes with additional increments, the value of such sums to less affluent would-be plaintiffs is greater than the value of such sums to more affluent would-be defendants. Insofar as the relevant sums of money are more useful to the less affluent would-be plaintiffs than to the more affluent would-be defendants, the disutilities associated with enabling these suits would, other things being equal, be outweighed by the utilities.

But some of these increases probably cannot be justified from a utilitarian standpoint. To the extent that equalizing access is likely to increase the number of meritorious lawsuits by less affluent plaintiffs, it is also likely to increase the number of non-meritorious lawsuits by such plaintiffs. As a general matter, such lawsuits tend to decrease overall utility—even apart from the costs of such litigation to taxpayers: while the losing plaintiff’s utility is at best unchanged by the loss, the winning defendant’s utility is diminished by having to expend material resources to defend against a lawsuit that lacked an adequate basis in law. Since non-meritorious lawsuits are rarely conducive to the utility of either plaintiffs or defendants, increases in such lawsuits cannot be justified on utilitarian grounds. Accordingly, to the extent that implementing SPEA would tend to increase the number of non-meritorious lawsuits, it would diminish net community utility.

There are, of course, a host of additional costs associated with devoting more resources to the civil justice system. Many citizens believe that the existing tax burden is unfair and resent being taxed for...
Towards a Theory of Legitimate Access

measures that effectively redistribute income. While one might reasonably believe that such resentment is misplaced, it is irrelevant from the standpoint of utilitarianism. What matters from this standpoint is only that net aggregate community utility is diminished and the needed tax increases are likely to significantly increase such resentment, thereby significantly diminishing public utility. Accordingly, insofar as the implementation of SPEA would impose burdens on taxpayers that diminish their utility (whether measured in terms of satisfied preferences, happiness, or pleasure), it results in further effects that decrease net aggregate community utility.

The general thrust of these admittedly speculative considerations, then, seems to point in the direction of a principle that, on the one hand, provides less affluent citizens with more access than they currently have but that, on the other, stops well short of providing perfectly equal access. While the benefits associated with an increase in meritorious lawsuits provide a plausible utilitarian argument for increasing the access of the poor to the civil justice system, the costs associated with an increase in non-meritorious lawsuits provide a plausible utilitarian argument for providing somewhat less than the perfectly equal access required by SPEA. If this is correct, then SPEA cannot be justified on utilitarian grounds.

C. The Rawlsian Conception of Justice as Fairness

Perhaps the most fundamental idea in John Rawls’s famous theory of justice as fairness is “the idea of a society as a fair system of social cooperation over time from one generation to the next.” Implicit in the claim that society is a fair system of cooperation, as Rawls understands that claim, are two further claims: (1) the terms that govern societal cooperation ought to be reasonably acceptable to each participant; and (2) those terms ought to be reasonable from the standpoint of the participant’s own prudential interests. Accordingly, Rawls attempts to

59. See Steve Farkas & Jean Johnson, Public Agenda, The Values We Live By: What Americans Want from Welfare Reform (1996) (indicating that sixty percent of whites and forty-eight percent of blacks believe that, at most, welfare should temporarily provide the “basics” in emergency circumstances).
60. It is important to reiterate that utilitarian arguments of this sort are inevitably speculative for the reasons discussed above in this subpart.
61. Rawls, supra note 30, at 5.
62. Id. at 6.
identify the fair terms of cooperation by means of a hypothetical agreement among rational participants: in his view, the principles of justice constraining the state’s lawmaking activities are those that would be chosen by rational persons in an “original position.”

The crucial idea of the original position is defined by three elements of normative theoretical importance. First, persons in the original position must be free and equal so as to preclude any unfair bargaining advantages among the parties. As Rawls points out, “if it is to be a valid agreement from the point of view of political justice[,] . . . these conditions must situate free and equal persons fairly and must not permit some to have unfair bargaining advantages over others.”

Second, persons in the original position are assumed to be concerned only to maximize their own interests, and are not assumed to take an interest in the welfare of other persons. The reason is that the most that can be assumed about the motivations of human beings is that they are motivated by their own prudential interests. While many humans are motivated by altruistic considerations, not all are. To ensure that the principles chosen by persons in the original position are universally acceptable, Rawls defines the original position in such a way that the only psychological assumptions on which it depends are true of every human being.

Third, and most importantly, persons in the original position are shielded from information about their own contingent abilities and circumstances by the so-called veil of ignorance. Persons behind the veil of ignorance do not know, for example, how smart, athletic, physically attractive, socially adept, wealthy, or healthy they are. As Rawls describes this feature of the original position:

[N]o one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like. Nor again does anyone know his conception of the good, the particulars of his rational plan of life, or even the special features of his psychology such as his aversion to risk or liability to optimism or pessimism. More than this, I assume that the parties do not know the particular circumstances of their own society.

---

63. RAWLS, supra note 19, at 15–19.
64. RAWLS, supra note 30, at 15.
65. RAWLS, supra note 19, at 10.
66. Id. at 118.
Towards a Theory of Legitimate Access

That is, they do not know its economic or political situation, or the level of civilization and culture it has been able to achieve. The persons in the original position have no information as to which generation they belong.67

Persons in the original position, then, know nothing about the abilities and properties that distinguish them from other persons. In effect, such persons know no more about themselves than they do about any other person; what knowledge they have about themselves is limited to knowledge of those properties that they share with every other person.

The point of the veil of ignorance is to seal off information that is irrelevant as far as justice is concerned.68 Although the principles of justice are chosen by rational agents concerned only to advance their own interests, they must do so only on the basis of information that is morally relevant. Information about people’s intellectual abilities is morally irrelevant because those abilities depend largely on circumstances over which they have little control: who their parents are, where they were born, and how much education they have are largely matters of luck. While such fortuitous circumstances are, of course, relevant with respect to one’s prudential deliberations, they are irrelevant with respect to one’s moral deliberations—and the choice of principles of justice is ultimately a moral choice. Accordingly, persons in the original position must choose principles that will advance their interests no matter what abilities and propensities they turn out to have.

The imposition of the veil of ignorance prevents persons in the original position from adopting an interest-maximizing principle for pursuing their prudential interests. In conditions of full information, rationally self-interested agents can pursue a strategy that aims at

67. Id.

68. As Rawls puts this important point:

[It] seems reasonable and generally acceptable that no one should be advantaged or disadvantaged by natural fortune or social circumstances in the choice of principles. It also seems widely agreed that it should be impossible to tailor principles to the circumstances of one’s own case. We should insure further that particular inclinations and aspirations, and persons’ conceptions of their good do not affect the principles adopted. The aim is to rule out those principles that it would be rational to propose for acceptance, however little the chance of success, only if one knew certain things that are irrelevant from the standpoint of justice. For example, if a man knew that he was wealthy, he might find it rational to advance the principle that various taxes for welfare measures be counted unjust; if he knew that he was poor, he would most likely propose the contrary principle.

Id. at 16–17.
maximizing their own utility. In particular, such agents can assess the expected value of each act $A$ by calculating the differential between the expected benefit of $A$ (i.e., the magnitude of the benefit associated with $A$ multiplied by its probability) and the expected cost of $A$ (i.e., the magnitude of the cost associated with $A$ multiplied by its probability) and selecting the act with the highest expected value.\(^{69}\) By selecting the act with the highest expected value, agents optimize their prospects for maximizing their own utility.

In conditions of highly restricted information, however, rationally self-interested agents must adopt a more conservative “maximin” strategy and choose behaviors that are minimally necessary to protect themselves against highly undesirable outcomes. As Rawls describes it, the maximin strategy “tells us to identify the worst outcome of each available alternative and then to adopt the alternative whose worst outcome is better than the worst outcomes of all the other alternatives.”\(^{70}\) The maximin rule, unlike the ordinary prudential strategy of maximizing expected value, takes into account only the relative magnitude of the worst possible outcomes; it does not take into account any information that assesses the comparative probabilities of the various options because such information is not available. In effect, then, rationally self-interested agents deploy the maximin strategy as a means for avoiding the worst of undesirable outcomes.

While some authors argue that the maximin strategy is not the only rational strategy applicable in situations of high risk and uncertainty,\(^{71}\) it should be clear that the maximin strategy is rationally deployed in such situations. A somewhat perverse example is helpful in illustrating the point. From the standpoint of prudential rationality alone,\(^{72}\) it is rational

---

\(^{69}\) While agents do not typically ground their behavior in explicit mathematical calculations of probability, this is true for a variety of reasons that do not call the general principle into question. First, in most instances, the relevant probabilities are known to be 1. For example, the probability of the only material cost of a candy bar—its price—is 1. Second, in circumstances in which the material probabilities are not known to be 1, the agent has only a rough feel for the relevant values. In such situations, the role that explicit calculations normally play is played by a rough intuitive process of weighing the outcomes.

\(^{70}\) RAWLS, supra note 30, at 97.


\(^{72}\) The standpoint of prudential rationality seeks nothing more than the maximization of one’s own interests. While it is probably true that a purely prudential standpoint is rarely appropriate, it is crucial to realize here that moral considerations are not relevant with respect to making purely prudential decisions. Accordingly, the following example brackets any considerations of morality,
Towards a Theory of Legitimate Access

for someone with full information to play the most dangerous games if the prize is large enough and the odds of losing are remote enough. Whether it is prudentially rational, for example, to play a game of Russian roulette with one bullet depends on the amount of the prize and on the number of empty chambers in the gun. While it would clearly be irrational to play if I know the prize is one dollar and there is only one empty chamber in the gun, it is clearly rational to play if I know the prize is one hundred million dollars and there are six billion empty chambers,73 one incurs a substantially greater risk of death every time one gets into an automobile. In these cases, there is sufficient information to adopt an interest-maximizing strategy that will sometimes dictate playing the game. But a more conservative maximin strategy is appropriate from the standpoint of prudential rationality if I lack information about some of the salient probabilities. For example, if I am not told how many empty chambers there are in the gun, it is clearly rational to adopt a maximin strategy that requires me to decline the game as a means of avoiding the worst of undesirable outcomes.

Although the motivation for the veil of ignorance is largely moral, its effect on the deliberations of agents in the original position is prudential in character. Since the veil of ignorance denies people any information about themselves that would tell them how likely they are to win or lose in society, they must adopt a more conservative prudential strategy for selecting the principles of justice than the interest-maximizing strategy that is available in conditions of full information. They must, as a matter of prudential rationality, choose those principles that are minimally necessary to enable them to avoid the very worst outcomes. Since a maximin strategy will enable them to do this, it is rationally deployed by agents in the original position.

Rawls believes that persons in the original position will avoid the very worst outcomes by choosing a principle that affords them maximum liberty compatible with everyone else’s having comparable liberty and a principle that assures that economic inequalities will conduce to their benefit no matter where they wind up in society. According to the

which would function to deter the agent from playing the game. The only issue in deciding whether or not to play is whether doing so maximally conduces to one’s self-interest.

73. This is subject to one qualification: if one believes that God exists and punishes suicide with eternal and infinite suffering in hell, then the expected cost of the game is infinite no matter how small the probability of losing. Since a finite expected benefit is infinitesimally small relative to an infinite expected cost, it is irrational to play the game on such theistic assumptions no matter how big the prize and how small the probability of losing.
Liberty Principle, “each person is to have an equal right to the most extensive scheme of equal basic liberties compatible with similar liberties for others.”74 According to the Difference Principle, “social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone’s advantage, and (b) attached to positions and offices open to all.”75 Persons in the original position thus use a maximin strategy to enable them to avoid catastrophic situations in which their freedom is denied or in which economic inequalities are permitted at their expense.

To evaluate SPEA from the standpoint of someone in the original position, then, we need to identify the worst outcomes of all the available alternatives to determine whether the worst outcomes under SPEA are preferable to the worst outcomes under all other principles. At the outset, it is easy to identify the very worst outcome: the very worst outcome is a catastrophic situation in which one has been severely injured by the negligent or willful behavior of another party and cannot enforce one’s right to compensation because one cannot access the civil justice system. In such circumstances, people have lost their capacity to make livings and lead meaningful lives because of the culpable behavior of another person, but cannot exercise their right to compensation from the former because they lack access to the civil justice system that would enforce their rights against the responsible party. Parties in this situation have no guarantee that they can even feed themselves; in a very literal sense, then, the very worst possible outcome is potentially life-threatening.

There are a number of principles that would permit this catastrophic outcome. This, for example, is the worst possible outcome of any principle that excludes, on its face, persons in any particular group from all access to the civil justice system for any reason whatsoever. But it is also the worst possible outcome of any principle that arbitrarily conditions access to the justice system upon the payment of a fee that is sizable enough to ensure that the least affluent people in a society cannot afford access to the justice system.76 To avoid this outcome, then, a person in the original position must select a principle that assures that no class of persons is excluded from access to the civil justice system either

74. RAWLS, supra note 19, at 53.
75. Id.
76. The reason for the qualifier “arbitrarily” is that, as will be argued below, people in the original position would not accept an absolute principle that requires the public to subsidize the access of the least affluent citizens because they have to guard against a situation in which they must sacrifice important material needs to subsidize another person’s access. See infra this Part & Part III.C.
Towards a Theory of Legitimate Access

for reasons of class membership or for financial reasons not related to the costs of such access and the quantity of material resources generally available to the society.

SPEA clearly satisfies this criterion. Insofar as SPEA assures that every person has perfectly equal access to the civil justice system, it is logically inconsistent with any principle that excludes, on its face, persons in any particular group from access to justice for whatever reason. Further, insofar as SPEA guarantees minimal access to the justice system, it is logically inconsistent with any principle that conditions access to the justice system upon payment of a fee sizable enough to permit some, but not all, to access the justice system.

While selecting SPEA would clearly enable an agent to avoid the very worst outcomes, SPEA is nonetheless problematic from the standpoint of someone in the original position. The worst outcomes under SPEA, though not as bad as the worst outcomes of a principle that has the effect of arbitrarily denying some persons access to the civil justice system, are sufficiently undesirable that a rationally self-interested agent would seek to avoid them from the original position. Because it is clear that the costs of perfectly equal access are substantial, they will have to be borne by comparatively affluent citizens. Accordingly, it is conceivable under SPEA that more affluent citizens will have to make significant material sacrifices to pay their share of the very substantial costs of perfectly equal access. While this will probably not require a person to defer satisfaction of a basic material need, it may require a person to defer or even give up satisfaction of a want that is sufficiently urgent or passionate as to provoke considerable unhappiness and resentment in the agent.77

In this connection, it is helpful to observe that although persons in the original position have some understanding of political and economic theory,78 they “do not know the particular circumstances of their own

77. Such a reaction would, for example, be quite natural in someone who sees a similarity between taxation for redistributive purposes and forced labor. As Robert Nozick expresses the similarity:

Seizing the results of someone’s labor is equivalent to seizing hours from him and directing him to carry on various activities . . . . This process whereby they take this decision from you makes them a part-owner of you; it gives them a property right in you. Just as having such partial control and power of decision, by right, over an animal or inanimate object would be to have a property right in it.

NOZICK, supra note 24, at 172.

78. RAWLS, supra note 19, at 119.
That means that they do not know whether the societies in which they live are comparatively prosperous or comparatively poor. Since they are thus shielded from any information that would tell them whether their tax burdens in subsidizing perfectly equal access will require significant sacrifices, the maximin strategy requires them to guard against the worst outcome—a state of affairs in which they have to sacrifice some of their most important wants to pay their share of the costs of perfectly equal access. Such prospects, then, provide rational self-interested agents in the original position with a powerful incentive to look for another principle that avoids such outcomes.

Moreover, it is clear that there are alternative principles that would avoid the very worst outcome of being denied access to the civil justice system. As it turns out, SPEA is not the least restrictive principle that would avoid these outcomes. Any principle that guarantees that people will not arbitrarily be denied access to the civil justice system will enable them to avoid the very worst outcome of being denied access to civil justice when it cannot be justified by considerations having to do with the costs of access and the quantity of resources available to society. Since such a principle is logically compatible with allowing some inequality of access, a person in the original position need not choose a principle that, in effect, guarantees that no person will have more convenient access to the civil justice system than any other person.

For these reasons, it is reasonable to think that SPEA would not be chosen by a rationally self-interested agent in the original position. Since the worst outcomes under SPEA involve significant disadvantage, a rationally self-interested agent would prefer any other principle that avoids the very worst possible outcome without allowing another outcome as disadvantageous as the worst outcome under SPEA. Further, since there are less burdensome principles than SPEA that avoid the very worst outcome, it is reasonable to think that someone adopting a maximin strategy would select one of these other principles. Accordingly, the adoption of SPEA cannot be justified by Rawls’s theory of justice.

79. Id. at 118.
80. See supra note 76 (explaining the qualifier “arbitrarily”).
Towards a Theory of Legitimate Access

III. TOWARDS A THEORY OF ACCESS: TWO PRINCIPLES

This part of the Article argues that the three different approaches to legitimacy converge on two principles that state necessary conditions for minimally legitimate access to the civil justice system. The first principle defines a negative obligation that requires only that the state refrain from certain types of action:

The Equality Principle (EP): The state is morally obligated not to exclude any person from access to any resource provided by the civil justice system for reasons that explicitly or implicitly deny the equal moral status of every human being.

As is readily evident, EP does not guarantee any positive right to access any particular resource; it requires only that the state not discriminate with respect to what resources it provides for reasons inconsistent with the moral equality of every human being.

Insofar as EP is therefore consistent with a state of affairs in which no person has access to any legal resources whatsoever, it operates in the following way. While EP does not require the state to provide access to any resources at all, it prohibits excluding people from access to any resource the state does provide for reasons that deny their status as moral persons. For example, while EP does not guarantee access to some resource to women, it implies that women may not be excluded from access to that resource for reasons that are sexist in character. To the extent, then, that any other class of human beings has access to that particular resource, EP requires that women must be afforded access to it unless there is some reason for excluding them that is not sexist in character.

By itself, satisfaction of EP could not be sufficient for morally legitimate access according to the adequacy condition described above at the end of Part I. As observed there, most classically liberal theories presuppose that the legitimizing function of the state is to promote the common good by serving the prudential interests of its citizens. Because this requires the state, at a minimum, to establish and protect legal rights in a manner that endows those rights with some prudential value to its citizens, no set of access principles can be sufficient for legitimate access without guaranteeing the prudential value of legal rights to the rights-holders. Because EP is logically compatible with a state of affairs in

81. As argued above, it is a human being’s status as a moral person that ensures his or her moral equality with every other human being. See supra Part I.
which no one has access to any legal resources and hence with a state of
affairs in which no one has the ability to activate the legal consequences
that endow legal rights with prudential value, EP could not exhaust the
conditions of morally legitimate access.

Accordingly, EP must be supplemented with a principle guaranteeing
that citizens have sufficient access to the civil justice system to endow
their legal rights with prudential value. What is needed, then, to
to adequately define a minimally legitimate right of access is a principle
that defines a positive right of access to what is minimally needed to
activate the relevant legal consequences. The following principle is an
affirmative norm that purports to do exactly that:

**The Reasonable Access Principle (RAP):** The state is obligated
to provide every person with reasonable access to all resources
that are minimally necessary (1) to develop an informed
plausible view about their legal rights and responsibilities and
(2) to competently prosecute or defend lawsuits implicating their
rights or responsibilities.

The operative concept of reasonableness in RAP is intended to govern
the distribution of costs among citizens and should be understood as
incorporating the following principles. First, the costs to any one
citizen of subsidizing another citizen’s access should not be greater than
the benefit to the latter and should not, in any case, require the former to
sacrifice satisfaction of needs or wants vital to well-being. Second,
subsidized access is reasonable only to the extent that its aggregate
public benefits exceed its aggregate public costs. Third, citizens should
not be forced by their own economic circumstances to forgo access to
minimally necessary legal resources when they have vital material
interests at stake in being able to defend legal positions unless access to
these resources cannot be subsidized consistently with the first two
principles. In effect, the first two principles define limits on the extent to
which the state can coercively require more affluent citizens to subsidize
the access of less affluent citizens to the civil justice system; the last
principle defines an affirmative obligation to transfer resources, subject

---

82. This assumes, of course, that the relevant legal consequences have prudential value. If the
legal consequences lack prudential value, then the ability to activate those consequences also lacks
prudential value.

83. What is and is not “reasonable” is determined by recourse to substantive normative standards.
As the term is normally used, the relevant standards include moral principles of fairness. To say, for
example, that people are behaving in a manner that is unreasonable, on this ordinary usage, is partly
to say that they are behaving unfairly.
Towards a Theory of Legitimate Access

to the first two principles, to ensure that the poorest citizens can defend vital material interests.84

Thus construed, RAP is sufficient to endow legal rights with positive prudential value to all rights-holders and hence satisfies the relevant adequacy condition. Regardless of how costs are distributed among the citizenry, RAP defines an unconditional obligation on the part of the state to provide every citizen with access to every resource minimally necessary to defend a plausible legal position. Accordingly, RAP defines a positive right on the part of every citizen to have access to resources that, taken together, are sufficient to enable a competent adult to defend an informed legal position.85 By itself, this assures that rational self-interested agents are able, in principle, to activate the legal consequences that endow legal rights with prudential value and hence, together with the prudential value of those consequences, guarantees the prudential value of legal rights.

Even so, it is important to emphasize that EP and RAP do not purport to exhaust the state’s obligations with respect to access. The claim here is that satisfaction of EP and RAP is necessary for a state’s civil justice practices to be legitimate, and not that satisfaction of EP and RAP is sufficient for a state’s civil justice practices to be legitimate. Strictly speaking, EP and RAP are compatible with the existence of additional principles that must be satisfied by the state for its civil justice practices to be legitimate—though, as we have seen, such principles will not include SPEA.86 In this sense, EP and RAP represent just a step towards

84. The rationale for defining “reasonable” in terms of whether certain benefits exceed certain costs is that the ordinary sense of the term incorporates substantive moral content and is hence incompatible with both utilitarianism and Rawls’s theory. It is incompatible with utilitarianism because utilitarianism assumes that all moral qualities of an act are fully determined by its effects on community utility. It is incompatible with Rawls’s view because persons in the original position are screened from their own particular moral views and are thus not in a position to agree on any principle that incorporates moral content. But while the operative notion is defined in purely non-normative terms, we will see that it is morally justified on each of the three conceptions of legitimacy. Thus, although the notion does not explicitly incorporate any substantive moral ideals, I will argue below that it satisfies three different sets of them. See infra Part IIIA–C.

85. Taken together, all the resources that are minimally necessary to develop and defend a legal position are minimally sufficient to do so. People who have access to every resource that is minimally necessary have access to everything they need in order to develop and defend legal positions; it follows that they have access to materials that are minimally sufficient to develop and defend legal positions.

86. Indeed, it is reasonable to think that the state is morally obligated to provide more resources to persons who are not sufficiently competent to defend their own interests; as is readily evident, EP and RAP are utterly indifferent with respect to such requirements.
a general theory of morally legitimate access, and not a fully comprehensive theory of legitimate access.

A. The General Principle of Equality

EP is a straightforward corollary of the general principle of equality. As will be recalled, the general principle of equality affirms the moral equality of all human beings \textit{qua} persons and hence requires that every human being be treated with equal respect—an obligation that applies to both persons and states. It is clear, as a logical matter, that any behavior towards human beings for reasons that deny them equal status as moral persons is inconsistent with the obligation to treat them with equal respect. The obligation to treat all persons with equal respect requires, at the very least, respect for their status as moral persons with full and complete sets of rights. To the extent, then, that the state is morally obligated to treat every human being with equal respect, it follows that the state may not exclude any person from any resource provided by the civil justice system for any reasons that explicitly or implicitly deny the equal moral status of every human being.\footnote{See supra Part II.A.}

While the principle of equality does not logically imply RAP, RAP coheres with the principle of equality and is hence justified by it. The obligation to treat each person with equal respect implies, \textit{a fortiori}, an obligation to treat each person with respect, and minimal moral respect for a person requires respect for his or her autonomy.\footnote{Most deontological theories of legitimacy treat autonomy as a basic value. John Rawls, for example, presupposes that agents in the original position are fully autonomous—a feature that is necessary to ensure the moral validity of the agreement on principles of justice. \textit{See} John Rawls, \textit{Political Liberalism} 77–81 (1993). In contrast, consequentialist theories protect autonomy only insofar as it conduces to the relevant favored state of affairs. Since the general principle of equality discussed here is a deontological one derived from deontological considerations about moral personhood, it is reasonable to infer that it requires minimal respect for autonomy.} Insofar as moral personhood confers upon a person full membership in the moral community with a complete set of moral rights and responsibilities, it entitles that person to respect for at least those decisions that do not violate the rights of other persons. Indeed, it is hard to make sense of the idea that one could treat a person with morally minimal respect without a commitment not to interfere with those decisions and behaviors that do not in any way implicate the rights of others.\footnote{To say this is not, however, to presuppose any particular conception of what rights others have. Such rights may be quite expansive as political liberals believe or less expansive as political conservatives believe. Thus, for instance, this statement is agnostic with respect to whether}
Towards a Theory of Legitimate Access

Respect for autonomy clearly requires that the state provide people with access in principle to what is minimally needed to adequately defend their legal interests. The obligation to respect autonomy undoubtedly implies substantive limits on the state’s legitimate lawmakering abilities; it presumably requires, at the very least, that the laws of the state allow persons a certain level of freedom to frame, express, and execute their own values as long as doing so does not violate the rights of others. But if respect for autonomy requires allowing people to steer their own course, it also requires, a fortiori, allowing people to steer their own course with respect to developing and defending informed and plausible legal positions. Since it is, by definition, impossible for people to develop and defend informed and plausible legal positions without having access to what is minimally necessary to do so, the state is obligated to allow, as a matter of principle, each citizen access to those resources that are necessary (1) to develop informed plausible views about their legal rights and responsibilities or (2) to prosecute or defend lawsuits implicating their rights or responsibilities.

It is, however, not enough that the state allows such access in principle; it must also take morally reasonable steps to ensure that every citizen can, as a practical matter, access the justice system when needed to defend legal interests that are vital to well-being. If the state is going to limit the options available to citizens for defending their vital interests by coercively enforcing a monopoly over the use of force in resolving disputes, then it must ensure that each citizen has morally reasonable access to what is minimally needed to defend his or her vital interests without resorting to force. It would be unfair for the state to preempt citizens’ ability to coercively defend their own interests without ensuring that they have morally reasonable opportunities to avail themselves of the state’s civil justice system. The obligation of equal respect seems to imply that the state take morally reasonable steps to ensure that all citizens are able to access the civil justice system (and hence the state’s monopoly on coercive force) when their vital interests are at stake.

consenting adults have an unrestricted right to engage in sexual relations—a subject of dispute between political conservatives and political liberals.

90. Because we are talking about what is required by deontological principles, the appropriate notion of reasonableness is the ordinary one that incorporates moral principles of fairness. I will argue that the principles that define the notion as it appears in RAP are reasonable in the ordinary moral sense of the word.
But given that the costs of providing access to such resources must ultimately be borne by citizens, there are limits to how much the state can, as a moral matter, reasonably demand from one class of citizens by way of subsidizing the access of another class of citizens. Equal respect for those whose incomes will have to be coercively taxed to subsidize such access surely entails limits on how much they may be taxed for such purposes; it seems clear, for example, that citizens could not legitimately be required to pay ninety percent of their incomes to assure some utopian level of access to the civil justice system.

If it is not entirely clear exactly where to draw the line on how much citizens can legitimately be taxed to ensure that others have adequate access, RAP certainly expresses a plausible candidate for such a limit. Equal respect for each citizen seems to require that no citizen be forced to absorb a cost that is greater than the benefit it makes possible to another citizen. Respect for citizens’ autonomy seems to mandate that they not be asked to subsidize other persons’ access if it requires them to sacrifice interests vital to their own well-being. Finally, respect for citizen autonomy seems to require recognition of the sacrifices that must be made to earn a living; coercively taxing citizens beyond the point where the public benefits exceed the public costs seems to squander those citizens’ hard-earned resources and hence seems inconsistent with any principle that requires recognition of the sacrifices they must make to earn a living. If this is correct, then RAP is, at the very least, presumptively justified under the general principle of equality.

B. Utilitarian Theories of Legitimacy

Utilitarian moral theories, including utilitarian theories of legitimacy, presuppose a principle of equal respect: according to this principle, no person’s utility—whether measured in terms of pleasure, happiness, or well-being—may be accorded greater weight in moral deliberations than any other person’s interests. As Peter Singer phrases this important point:

[W]hen I make an ethical judgment [under utilitarian moral theories,] I must go beyond a personal or sectional point of view and take into account the interests of all those affected [by our actions]. This means that we weigh up interests, considered simply as interests and not as my interests, or the interests of

91. Again the relevant sense of “reasonable” here is the ordinary sense. See supra note 90.
Towards a Theory of Legitimate Access

Australians, or of people of European descent. This provides us with a basic principle of equality: the principle of equal consideration . . . [requires] that we give equal weight in our moral deliberations to the like interests of all those affected by our actions.92 Whatever it is, then, that confers membership in the community of moral persons according to utilitarianism,93 it requires that every moral agent, individual or institutional, accord each person equal respect by assigning as much weight to that person’s utility as to any other person’s utility.94 Interpreted in light of this foundational commitment, utilitarian theories of legitimacy clearly imply EP. The argument is straightforward. Insofar as the state excludes any person from access to some legal resource for a reason that denies the equal moral status of every human being, it disregards the utility of that person relative to those of other persons. To utterly exclude someone from a resource for such reasons is to do so without any regard for whether access to that resource would conduce to her utility. But disregarding the utility of one person P in making a moral decision about her is equivalent in effect to a strong form of discounting P’s utility relative to the utility of other persons; it has exactly the same effect as always assigning no value to P’s utility because it makes decisions about how to treat P turn entirely on the consequences to the utility of other persons.95 Since the state is therefore obligated not to disregard the utility of any person under utilitarianism and since the state cannot exclude a person from access to legal resources for any reason that denies her equality without utter disregard for that person’s utility, it follows that the state is obligated not to disregard a person’s utility for a reason that denies her equality.

92. PETER SINGER, PRACTICAL ETHICS 21 (2d ed. 1993).
93. While this is sometimes thought to involve the capacity for rationality, Peter Singer argues that the capacity for suffering confers a right to equal consideration on animals. See id. ch. 2.
94. Indeed, this feature of utilitarianism has frequently been criticized as setting a standard that is unreasonably high. For example, Bernard Williams argues that the duty of equal consideration requires persons to subordinate the projects that give meaning to their lives in order to promote the utility of other persons and is hence inconsistent with respect for autonomy. Thus, for example, utilitarianism would require me to give up philosophy, which I love, to study medicine, which I do not, if doing so would maximally promote utility. See J.J.C. SMART & BERNARD WILLIAMS, UTILITARIANISM: FOR AND AGAINST (1973).
95. For example, a state that denies women the right to vote on the ground that women do not have the same moral status as men has effectively excluded women from voting for reasons that have nothing to do with the utility to women of being able to vote and everything to do with the utility to men of women not being allowed to vote.
Accordingly, if utilitarianism’s principle of equal consideration is true, then EP must also be true.

To justify RAP from the standpoint of utilitarianism, we need to go beyond utilitarianism’s formal commitment to equal respect and assess the relevant utilities and disutilities. Utilitarianism, as will be recalled, defines the moral value of an act entirely in terms of its extrinsic effects on community well-being. Thus, utilitarian theories of legitimacy imply that the sole obligation of the state is to act in ways that maximally promote net aggregate utility among its citizens. While EP can be justified as a logical consequence of utilitarianism’s principle of equal consideration, RAP is not a logical consequence of this principle because it is consistent with much stronger principles like SPEA. 96 For this reason, whether RAP is justified under utilitarianism depends entirely on its effects on net aggregate community utility.

RAP’s guarantee of access to resources that suffice to enable a competent adult to defend a legal position promotes the utility of potential plaintiffs. 97 Rational agents would experience significant disutility if seriously injured by the culpable conduct of someone they could not sue because they lacked access to the necessary legal resources. It is important to note, however, that the relevant disutilities are not limited to the obvious ones associated with not being able to extract compensation from the culpable party, which are greater for less affluent would-be plaintiffs than for more affluent would-be plaintiffs. They also include the resentment that naturally accompanies being denied access to the civil justice system when needed to hold a culpable party accountable for serious injuries, which are presumably as significant for more affluent would-be plaintiffs as they are for less affluent would-be plaintiffs.

For similar reasons, RAP’s guarantee of access generally promotes the utility of potential defendants. Rational agents would incur significant disutility if they were defendants in a civil lawsuit seeking substantial money damages without having access to the resources

---

96. It is a basic theorem of logic that if each of two inconsistent propositions is consistent with a proposition $A$, then $A$ implies neither of those propositions. For example, the proposition that apples are red is inconsistent with the proposition that George Bush is president and with the proposition that George Bush is not president. Hence the proposition that apples are red does not imply either of those latter propositions. See, e.g., ELLIOT MENDELSON, INTRODUCTION TO MATHEMATICAL LOGIC (4th ed. 1997).

97. It should be recalled here that access to every resource minimally necessary for such purposes implies access to resources that, taken together, are minimally sufficient for such purposes. See supra note 85.
Towards a Theory of Legitimate Access

needed to pursue plausible legal defenses. As before, the relevant disutilities are not limited to those associated with the unfavorable outcome—though having to pay substantial money damages is undeniably a very unhappy outcome for most people. They also include the significant unhappiness most people would experience at the very prospect of not being able to tell their sides of the story because they are denied access to the civil justice system.98

Nevertheless, a positive right of access also creates theoretically significant disutilities. Insofar as RAP requires that people have access to every resource minimally necessary to defend an informed legal position, it requires that the state incur significant expenses that must be passed on to taxpayers. As one might expect, the costs to the state of creating, sustaining, and making accessible resources it would not otherwise have to worry about are substantial. A recent survey showed that the direct costs of tort litigation in the United States reached $205 billion in 2001—2.04% of the gross domestic product that year;99 given that tort disputes are only one source of civil litigation, one can expect that the costs are significantly higher for civil litigation as a whole. But insofar as the costs of such resources have to be borne by taxpayers, the provision of those very resources impose disutilities on the very persons who are supposed to benefit from those resources; indeed, the direct costs to each U.S. citizen of tort litigation was $721 in that year,100 which amounted to a “litigation tax” of almost $2900 for a family of four. As is readily evident, these costs impose significant disutilities on the very citizens who stand to benefit from RAP.

Even so, it seems clear that the utilities outweigh the disutilities for most people—at least if their behavior in other contexts is any indication. Most people, for example, experience considerable unhappiness at even the prospect of developing a serious illness without being able to afford appropriate medical care. To eliminate this prospect, people who can afford it typically purchase health insurance despite the substantial disutility it imposes in the form of expensive monthly premiums. While it is unfortunately true that there are too many

98. This class of disutilities helps to explain why rights matter so much to us: they give us peace of mind by purporting to diminish (if not eliminate) the prospect that some evil will be done to us.


100. Id. at 11.
Americans without adequate health insurance, it is because they cannot afford it—and not because they do not want it. Rational self-interested persons are typically willing to shift material resources, when they can afford to do so, from the pursuit of other material goods to purchase insurance policies that protect them against the risk of incurring very significant disutilities. Indeed, this is why insurance companies are typically profitable to begin with.

The same sort of reasoning applies to RAP. While the provision of such resources must be subsidized by taxpayers at considerable expense to themselves, it is reasonable to hypothesize that taxpayers regard the disutility associated with the costs of such measures as being significantly outweighed by the utility that comes with the peace of mind of knowing that one will not find oneself in a situation in which one needs to develop and defend a plausible legal position but cannot do so because one is denied access to what is minimally needed to do so. RAP can thus be justified under utilitarian theories of legitimacy in the same way that the purchase of health insurance can be justified by comparison of the relevant utilities and disutilities associated with buying and not buying it.

Of course, there is one difference between the two situations. Whereas individuals are typically responsible for only the costs of their own health insurance, RAP will sometimes make some individuals responsible for the costs of ensuring effective access of other individuals. Insofar as RAP authorizes the state to require that one class of citizens subsidize the access of another class of citizens, it contains a redistributive element that could arguably give rise to significant disutility in the form of resentment among those who must subsidize the access of other persons.

Nevertheless, RAP minimizes the risk that such disutilities would ultimately outweigh the utilities of subsidized access by sharply limiting the extent to which any citizen can be asked to subsidize the access of less affluent citizens. For starters, RAP does not permit the state to require one citizen to subsidize another’s access when doing so would force the citizen to sacrifice the satisfaction of important material wants. In addition, RAP attempts to ensure that transfers of wealth are cost-efficient in two important respects—one individual and the other

Towards a Theory of Legitimate Access

collective. First, RAP ensures that an individual will never be asked to bear a personal cost that is out of proportion to the benefit to others. Second, RAP ensures that the community as a whole will not be taxed for measures that do not result in aggregate public benefits that exceed their aggregate public costs. These features of RAP diminish the risk of significant citizen resentment by reducing the likelihood that citizens will regard their responsibility for subsidizing other citizens’ access as unfair or excessive—natural sources of the sort of resentment that might ultimately undermine RAP’s justifiability under utilitarian principles. For such reasons, it is plausible to conclude that RAP is justified under utilitarian theories of legitimacy.

Even so, the possibility of such public resentment calls attention to one potentially troubling feature of RAP. Insofar as consideration of public costs includes such subjective factors as taxpayer resentment, RAP implies that redistributive measures that guarantee effective access for less affluent citizens are justified only to the extent that they are minimally acceptable to the taxpaying public. While it would take considerable unhappiness on the part of taxpayers to outweigh the obvious utilities that such redistributive measures create, one can conceive of states of affairs in which taxpayer resentment exceeds the utility of such access to the poor. If, for example, taxpayers are sufficiently self-centered and beneficiaries are sufficiently disinterested, taxpayer resentment would outweigh the utilities to the poor of having effective access to the civil justice system. In that situation, RAP would not authorize the state to adopt redistributive measures to ensure that less affluent citizens can afford to access the civil justice system—though, of course, RAP would in such circumstances guarantee them access in principle.

Indeed, one might be tempted to think that this admittedly disturbing possibility constitutes a counterexample that refutes RAP, but this would be a mistake. As unfortunate as this state of affairs would admittedly be, it is compatible with other important shared moral commitments. For example, such a state of affairs is compatible with the general principle of equality and the underlying principles defining the status of moral personhood;102 as argued in the last subpart, respect for people’s autonomy limits what can be coercively required from them by way of material contribution to the well-being of other persons. If a comparison

102. As we will see in the next subpart, it also harmonizes nicely with Rawls’s conception of justice as fairness.
of relevant benefits and costs is not the ultimate standard of morality, it certainly limits the extent to which the resources of one person can legitimately be transferred by coercive measures like taxation to promote the utility of other persons. For this reason, it would arguably be wrong for the state to coercively require one set of persons to subsidize access for other persons if the public costs exceeded the public benefits. Thus, the worst-case scenario under utilitarianism is not a counterexample that warrants rejecting RAP.

C. The Rawlsian Conception of Justice as Fairness

As will be recalled, the principles of justice defining the state’s legitimate lawmaking authority are, on Rawls’s view, those that would be chosen from a position of highly limited information by free and rational persons concerned only to advance their own interests. The Rawlsian veil of ignorance deprives people of any information that would enable them to assign probabilities to the various ways in which their lives might turn out on the ground that such information has nothing to do with what is just or fair. While the rationale for depriving people of such information is moral in character, the effect of the veil of ignorance is practical: it forces people to move from an interest-maximizing strategy to a more conservative maximin strategy that looks to avoid the very worst outcomes.

In evaluating whether EP and RAP would be chosen from the original position, it is crucial to emphasize the role that minimal access plays in making possible the development and pursuit of a plausible legal position. People who lack access to what is minimally needed—and hence necessary—for pursuing plausible legal positions cannot even begin to protect their interests in the case of legal disputes: people cannot competently prosecute or defend lawsuits if they do not have access to what is necessary to do so. Indeed, it is true, as a logical matter, that it is

103. Here it is helpful to note that this unfortunate state of affairs is only likely to occur in extremely poor societies in which most persons struggle to meet their basic needs. In such societies, the state would not be obligated to shift resources from some persons to pay for the access of other persons because this could not be done without requiring the former to sacrifice satisfaction of material wants—something that no one can legitimately be coerced to do. In somewhat more affluent societies where most persons have a small surplus over what is needed to satisfy their basic needs, RAP would require, at most, shifting resources to assure that every person has access to sufficient legal resources in cases where vital basic interests are at stake. From the standpoint of ordinary moral commitments, these results seem quite reasonable.

104. See supra Part II.C.
Towards a Theory of Legitimate Access

not possible for a person to develop or defend an informed legal position without access to what is necessary to do so.\textsuperscript{105}

The value of such ability to rational agents concerned to advance only their own interests from an original position of limited information is clear. As far as principles of access are concerned, the worst possible outcomes for rational self-interested agents worried about the possibility of legal conflicts with other people involve situations in which they are denied access, whether for legal or factual reasons, to what is minimally needed to defend their legal interests. It is true, of course, that how bad these outcomes are for agents varies indirectly with their material resources; other things being equal, the more material resources they have, the better they can tolerate the costs of being denied access to what they need to defend their interests. But since, from the original position, the agents do not know what their material prospects are, they must seek to avoid the worst outcomes. And, from the standpoint of people who lack information about their own material prospects, the very worst outcome is utterly unacceptable: to be completely denied access to what is minimally necessary to defend important legal interests. People in such situations face substantial threats to their well-being.

Accordingly, rational self-interested agents employing a maximin strategy from the original position would reject any principle of access compatible with such outcomes. As will be recalled, the maximin strategy requires rational self-interested agents to minimize the worst possible outcome by choosing the principle that allows for the least objectionable among worst possible outcomes.\textsuperscript{106} Thus, if there are any principles with a less objectionable worst outcome than complete lack of access, a person in the original position would choose from among such principles. From the standpoint of prudential rationality, any principle that is compatible with being utterly denied access to what is minimally needed to defend a legal position is unacceptable.

The maximin strategy dictates the adoption of EP because any principle that logically implies the negation of EP (i.e., the general “principle” that it is permissible to deny human beings access to legal resources for reasons incompatible with their status as moral persons) is compatible with these unacceptable outcomes. Insofar as the state adopts a rule that allows it, as a general matter, to deny human beings access to

\textsuperscript{105} The proposition that it is necessary that \( P \) is logically equivalent to it is not possible that not-\( P \). See, e.g., Brian F. Chellas, Modal Logic: An Introduction (1980).

\textsuperscript{106} See supra Part II.C.
legal resources for these reasons, it creates the possibility that agents in the original position might be utterly denied access to what they need to defend legal positions. Since the worst outcomes under EP involve, at most, (misplaced) resentment about someone’s access to the justice system, the maximin strategy requires rejecting any principle that implies the negation of EP (which, of course, includes the negation of EP itself). Accordingly, the maximin strategy requires the adoption of EP to assure agents that they cannot permissibly be denied access to something they need to defend their legal interests for reasons that deny their equality as moral persons.

Nevertheless, adopting EP is not sufficient to fully eliminate outcomes that are unacceptable from the original position. The reason for this is that EP, by itself, does not suffice to guarantee access of all persons to what is minimally needed to defend one’s legal interests when necessary. Insofar as EP prohibits only exclusion for discriminatory reasons (i.e., reasons that explicitly or implicitly deny any human being’s status as a moral person), it is logically compatible with a state of affairs in which a person is denied access for non-discriminatory reasons. As a logical matter, the state would not violate EP by requiring payment of prohibitively expensive fees for access to legal resources that do not reflect the costs as long as such a measure is not logically grounded in a denial of any human being’s personhood.

To block all such unacceptable outcomes, people employing a maximin strategy would have to begin by choosing a rule that guarantees them access in principle to what is needed to defend their legal interests—no matter what societal or generational circumstances in which they find themselves. From the standpoint of prudential rationality, the prospect of being arbitrarily denied access to what they need to defend a legal position when significant material interests are at stake is simply unacceptable. For this reason, rational self-interested agents employing a maximin strategy from the original position would choose a rule that, like RAP, defines a positive right that guarantees access, in principle, to the civil justice system.

But a principle that guarantees only access in principle will not suffice to block all outcomes in which the agent in the original position is

---

107. In particular, it would involve resentment that some human being who is mistakenly thought not to be a person has access to the justice system. It should be clear that this is a less objectionable outcome from the standpoint of prudential rationality than being denied access for reasons that deny one’s personhood.

108. Every proposition logically implies itself.
Towards a Theory of Legitimate Access

denied access to the legal system, because a person who is guaranteed access in principle might lack the needed financial resources to take advantage of such access. From the original position, however, the selection of a principle that minimizes the risk of these outcomes is complicated by the fact that principles assuring meaningfully effective access to the civil justice system will likely require coercive redistributive measures that implicate other interests that agents in the original position are concerned to protect. Accordingly, the agent seeking a principle that provides greater access than one guaranteeing access in principle will have to give some thought to how the costs will be distributed.

In deciding among principles of access, then, rational self-interested agents in the original position must guard against two outcomes that threaten vital interests. First, since the claim that one has access in principle does not imply the claim that one has adequate material resources to effectively access the civil justice system, agents must guard against the possibility of being unable to afford access to the civil justice system when they need it to defend vital legal interests. Second, because citizens will ultimately be coercively taxed to pay for the costs of access, agents must guard against the possibility of having to sacrifice significant material wants to subsidize someone else’s access to the justice system.

Unfortunately, there is no principle that would simultaneously foreclose the possibility of both outcomes. Any principle that forecloses the possibility of anyone’s being denied effective access to the justice system can do so only by leaving open the possibility that the agent might be required to forego satisfaction of significant material wants to subsidize the access of less affluent citizens. Depending on the economic situation of the society, a guarantee of effective access might require significant sacrifices on the part of more affluent citizens. Similarly, any principle that utterly forecloses the possibility of anyone’s having to forego satisfaction of significant material wants to subsidize someone else’s access creates the possibility that the agent might be denied access to the civil justice system for economic reasons. Depending on the economic situation of the society, a guarantee of limited taxation for subsidized access might have the effect that some less affluent citizens cannot afford to access the civil justice system when they need to.

Accordingly, it is reasonable to think that agents in the original position would attempt to achieve exactly the sort of balance between the two outcomes that RAP purports to achieve. To minimize the risk
that the agents might be required to make material sacrifices to subsidize someone else’s access that are not justified by the benefits, they would choose a principle that limits the extent to which they can be coercively required to bear the costs of someone else’s access: they will never be asked to make either significant material sacrifices or sacrifices that are greater than the benefits they make possible for other persons. To minimize the risk that the agents might be unable to access the civil justice system in cases where their vital interests are at stake, they would choose a principle that requires the state to guarantee effective access in such cases to the extent that it does not require those subsidizing such access to sacrifice their own vital interests.

For such reasons, agents in the original position would guard against such undesirable outcomes by choosing a principle that guarantees, at most, access to the justice system that is “reasonable” in the sense that the benefits to those who receive such access outweigh the costs to those who must subsidize such access and require no citizen to sacrifice vital interests to subsidize another citizen’s access. Agents pursuing a maximin strategy from the original position must find a prudentially acceptable way to optimize the worst outcomes to which they might be subject depending on how their lives go. They must therefore guard against outcomes that, on the one hand, deny them access to what they need to defend legal positions and, on the other, require them to sacrifice significant material desires to reduce the inconveniences of accessing justice when the public costs of doing so exceed the public benefits. If we construe the notion of “reasonableness” as being defined in terms of public benefits outweighing public costs, then RAP will optimize the worst possible outcome to which an agent in the original position might be subject. Thus, it is fair to conclude that an agent in the original position would supplement EP with RAP as a means of excluding all of the most unacceptable outcomes.

IV. CONCLUSION

This Article attempted to take a significant step towards a general theory of morally legitimate access. It argued that none of the three most influential approaches to morally legitimate authority requires the state to provide citizens with perfectly equal access to the civil justice system; it is legitimate for the state to allow some inequalities with respect to how easy or burdensome it is for a citizen to access the civil justice system. Further, the Article argued that these three approaches to legitimacy converge on two principles: one that defines an affirmative
Towards a Theory of Legitimate Access

obligation (i.e., the Reasonable Access Principle) to provide to each citizen what is minimally necessary to develop and defend a plausible legal position and one that defines a negative obligation (i.e., the Equality Principle) to refrain from restricting access to the civil justice system for reasons that deny the equality of every moral person. Taken together, these two principles define necessary conditions for the legitimacy of a state’s civil justice practice. While there is surely much to be done in working out a theory of legitimate access, these two principles represent a significant step towards a fully comprehensive theory of access.
TECHNOLOGY, VALUES, AND THE JUSTICE SYSTEM: 
THE EVOLUTION OF THE ACCESS TO JUSTICE TECHNOLOGY BILL OF RIGHTS* 

Donald J Horowitz†

I. GENESIS AND ROOTS

Almost four years ago, in early 2000, King County Law Librarian Jean Holcomb, a member of the Technology Committee of the Washington State Access to Justice (ATJ) Board, had an idea. She and other committee members had exchanged thoughts about how new information and communication technologies, including the Internet, were entering the justice system. They realized that the volume and speed of this flow would grow substantially and that it would not take very long before the new technologies would permeate the justice system.

Jean and the committee recognized that technology has the potential to provide increased pathways for access to justice, but that it can also do just the opposite—perpetuate and even worsen existing barriers, disparities, and exclusions, and in fact create significant new barriers, and exclusions. It was then that Jean decided to try out her idea. In May 2000, she wrote an article in the Washington State Bar News entitled The Digital Divide and Digital Justice: Do Clients Need a Technology Bill of Rights?1 In this article Jean demonstrated that the dissemination of ideas and information can have potent effects. This theme developed, and eventually resulted in what is today known as the Access to Justice Technology Bill of Rights (ATJ-TBoR).2

* See also the article by Richard Zorza in this volume, in which he describes and discusses the Access to Justice Technology Bill of Rights Project from a different perspective. Richard Zorza, Some Reflections on Long-Term Lessons and Implications of the Access to Justice Technology Bill of Rights Process, 79 Wash. L. Rev. 389 (2004). We worked closely together throughout the project, and I believe his observations and reflections are very valuable.
† Former King County Superior Court Judge; Chair, Access to Justice Technology Bill of Rights Committee.
2. Washington State Access to Justice Technology Principles (popularly referred to as the Access
The historical roots go back a few years earlier. On April 18, 1994, the Washington State Supreme Court entered an Order that would turn out to have a significant impact on justice in the state—and ultimately elsewhere. The Order established the Access to Justice (ATJ) Board and declared as its premise that “Washington State’s justice system is founded on the fundamental principle that the justice system is accessible to all persons.” The Court recognized that access to justice is a basic and accepted value, and that the effort to transform that aspiration into reality must be consistent and ongoing. The Court embraced the reality that access to justice is essential to a democratic system and to attaining its ultimate goal—equal justice for all.

To transform these values into reality, the Washington State Supreme Court gave the ATJ Board the mission to promote and facilitate equal access to justice, and, among other tasks, to develop and implement policies and initiatives that enhance, improve, and strengthen access to justice. On November 2, 2000, the Court entered an Order which reauthorized the ATJ Board as a permanent body, charging it with responsibility to assure high quality access for all persons in Washington State who suffer disparate access barriers to the justice system. The Court gave the ATJ Board the specific task, among others, to “develop and implement new programs and innovative measures designed to expand access to justice in Washington State.”

In every one of its Orders, the Washington State Supreme Court repeated its declaration that access to justice is a fundamental right of all people. These were beacons that many people in Washington State would soon begin to follow, but they were beacons in a somber context.

---

4. *Id*. This is also the very first statement in the “Access to Justice” section of the website of the Washington State Bar Association. See http://www.wsba.org/atj (last visited Dec. 27, 2003).
II. THE NEED FOR AN ACCESS TO JUSTICE TECHNOLOGY BILL OF RIGHTS

Historically, American society has been substantially divided between people who have access to essential resources—economic opportunity, health care, education, shelter, justice—and those who either experience great difficulties in securing access to such resources, or do not have access at all.

Despite its stated principles, the American justice system has not easily or quickly moved to eliminate existing barriers which impede or block access and opportunity for the poor, ethnic and racial minorities, persons with physical or cognitive disabilities or limitations, children and others easily taken advantage of or abused, or the elderly—indeed all who are vulnerable. As with the rest of society, the law has too often been an instrument of perpetuating social inequities, and in times of change has often adapted its old barriers and applied them to new situations. Again like the rest of society, when major innovations have presented themselves, the development, introduction, and use of these innovations has often mirrored and perpetuated existing inequities and frequently caused new disparities as well. Major opportunities for social progress and equity have been lost, due at least as much to lack of forethought and vision, lazy thinking, and just plain inertia as to malevolent intention.

As organizations, institutions, and systems transfer their information, operations, and a wide range of programs to computer-based information technology and use a variety of new communication technologies, they are also now creating what many call a “Digital Divide.” A central aspect of what is meant by the Digital Divide is the dynamic of technology-caused inequality. The Digital Divide is not so much a problem intrinsic to technology as a statement of how technology and lack of access to it can perpetuate and increase significant power imbalances among groups of people, and in fact create new imbalances. As societal systems increase their use of information, communication, and associated technologies to perform their basic functions of exchanging information and delivering essential services, opportunities, and products, many people who are already marginalized are finding technology to be a new barrier rather than a pathway to meeting their essential needs, creating greater disparity between the haves and have-nots, making things worse, not better.
We are now at a decisive point! The incontestable historical fact is that information and communications technologies do offer to all people, not just a privileged few, the greatest potential for access to information and consequent transformative opportunity since the invention of moveable type over four centuries ago. The historical challenge is that at the same time, these new technologies also threaten to isolate even further those who historically have and currently still experience significant barriers to economic opportunities, health care, shelter, education, and both social and legal justice. Thus, while significant opportunities and pathways leading toward equity and equality are visible, the existence of an additional disparity, a Digital Divide, superimposed upon centuries of pre-digital disparities, looms very large and threatens to eliminate these opportunities, block these pathways, and make the situation even worse than before.

III. THE ROLE OF THE ACCESS TO JUSTICE BOARD IN CREATING THE ACCESS TO JUSTICE TECHNOLOGY BILL OF RIGHTS INITIATIVE

During the late 1990s, the ATJ Board came to believe that the recent and ongoing developments in information, communication, and associated technologies, including the Internet, and the current and future use of such technologies pose significant challenges and significant opportunities to full and equal access to the justice system. Its members came to understand that technology can provide increased pathways for access to justice, but it can also perpetuate or worsen existing barriers and create significant new barriers. Given its mission, the ATJ Board became dedicated to ensuring that technology-generated barriers to accessing the justice system are avoided, eliminated, or minimized, and that opportunities and pathways to the justice system are created or maximized.

The ATJ Board understood that technological innovations and changes and their application and adoption were still in their early stages, and that as yet only a few waves had been felt. However, the ATJ Board also recognized that a great volume of change, indeed a transformation, was building that would inevitably and significantly impact access to and the quality of the system. With this in mind, the Board concluded that in the absence of careful deliberation, planning, preparation, and action, these enormous oncoming changes could have the destructive effects of a tsunami tidal wave. If, however, this great energy was prepared for and constructively channeled, the public and the
The Evolution of the ATJ-TBoR

justice system would not only avoid significant damage but would likely garner substantial benefits and create a more accessible, equitable, and effective system for all.

After a period of investigation to determine the right course, the ATJ Board determined that an Access to Justice Technology Bill of Rights premised on both federal and state constitutional precepts and our society’s core values should be developed and implemented. This body of fundamental principles would apply to all persons and groups, including but not limited to users and potential users of the justice system and those working in or in association with the justice system. This ATJ-TBoR initiative was created to accomplish the following tasks and goals:

1. Develop and implement an Access to Justice Technology Bill of Rights premised on relevant principles contained in the United States and Washington State Constitutions, the mission and underlying principles and declarations generating the creation and operation of the Access to Justice Board, the principles contained in the Hallmarks of an Effective Statewide Civil Legal Services Delivery System adopted by the Access to Justice Board in 1995, and subsequent and effectuating documents and declarations.7

2. Identify the strategies, means and methods to ensure that the rights and principles contained in the Access to Justice Technology Bill of Rights are adopted, become publicly known and accepted, and have concrete, practical and effective consequences in the daily lives of all people in the State of Washington.

The methods and goals of the ATJ-TBoR Committee were to:

1. Take optimal advantage of the unique opportunity provided by the confluence of time, place, resources, values, and will at this moment in history so as to increase both access to the justice system and the quality and equality of justice delivered to all persons and groups within our scope of service and influence.

2. Develop, declare, adopt, and implement a living body of just principles which in an ongoing way permeate and influence the justice system in the State of Washington and the lives and conduct of all persons or groups involved with or affected by the

justice system. To the extent appropriate and acceptable to other states, jurisdictions, and sectors throughout the United States and abroad, provide a model that may constructively be used or adapted.

3. Accomplish the foregoing in a manner that is thoughtful, balanced, and connected to the realities of life, with implementation that is practical, guides consequences, and takes into account those who provide the services in the system and the end user. In the course of so doing, listen to, inform and build a broad-based constituency; develop a public and political will and a collaborative momentum deeply committed to creating and maintaining access to and quality equal justice in the daily lives of all persons. 8

4. For the quality, credibility, and legitimacy of the process and the products, it is essential that the Committee and the process reaches out, receives, listens to, and in fact uses information, viewpoints, and suggestions from people and groups representing a broad array of backgrounds, experiences, perspectives, and expertise, never neglecting to include those the system is meant to serve—its consumers and end users. Inclusiveness is essential.

IV. THE ORGANIZING STEPS

The new ATJ-TBoR Committee of the Washington State ATJ Board held a major organizing meeting in May 2001. After soliciting and attracting a group of volunteers from various backgrounds, experiences, and disciplines who were willing to commit time and energy, the Committee and the initiative began its formal work in September 2001. First, the group developed a vision of what the effort was about, and then a process to achieve its goals and objectives in concrete form. The adopted Mission Statement was: “To create a body of enforceable fundamental principles to ensure that current and future technology both increases opportunities and eliminates barriers to access to and effective utilization of the justice system, thereby improving the quality of justice for all persons in Washington State.”

As the first and still the only such public policy initiative in the country, it quickly became apparent to those closely involved in the

---

The Evolution of the ATJ-TBoR

effort that this initiative was not directed solely at solving a justice system problem, but was at its core addressing fundamental issues of social justice and equity in a very broad sense.

Working committees were formed and filled by volunteers with knowledge, experience, or background in or relevant to the area of the committee or group focus. The project attracted well in excess of one hundred volunteers of diverse backgrounds and experiences who gave generously of their time and knowledge. There was only one paid full-time staff member, our coordinator, executive director, and versatile factotum at large. All funds for his salary and project expenses were privately raised, costing the ATJ Board, the Bar Association, and the public no money. 9

The ATJ-TBoR enterprise formally concludes in January 2004, by which time the Access to Justice Technology Principles, popularly known as the ATJ-TBoR, and their effectuating products and mechanisms will be well on their way to being adopted, institutionalized, and becoming no longer a special project but an intrinsic and ongoing part of the justice system and related systems. Projects and efforts underway will be completed, but no new initiatives will be undertaken by this Committee or its subgroups. Appropriate places will be found for new enterprises and tasks to be accomplished. The Steering Committee will continue to exist to supervise the completion of ongoing efforts, advise regarding the placement of new efforts, and aid in the process of institutionalizing and perpetuating the values and effectuation of the ATJ-TBoR process and its products and mechanisms.

V. THE ONGOING CHALLENGES OF TECHNOLOGY ADDRESSED

As the process moved forward, it came to focus increasingly on the concrete realities of the impact of technology on access to justice.

Like most of our society’s central institutions, the justice system is rapidly converting its information base and transaction systems to the new information and communications technologies. Some benefits of these new technologies are already obvious. Using a computer at home or at a nearby library branch or community center, people can initiate or respond to court or other legal requirements, communicate, and

---

9. We are grateful for grants awarded to the ATJ-TBoR initiative by the Open Society Institute, the Horowitz Foundation, the Legal Services Corporation, the Markle Foundation, the Paul Allen Foundation, and the State Justice Institute.
exchange documents with their lawyer or others associated with the legal system. In these cases, using technology means not having to travel to a court or office, which means less travel time, inconvenience and cost, less time away from work and family, less or no copying, mailing, or similar costs. Such ease can be especially important for the elderly, persons with disabilities, persons with limited incomes, and those who literally cannot afford to miss time from work, which can affect their income or even jeopardize their jobs. As informal library surveys are showing, many persons, including those with limited mobility or hearing, can seek and get information electronically about their rights as tenants; victims of domestic violence can learn on the Internet what they can do to protect themselves and can even start legal proceedings from a place they can get to more easily and safely than they can get to the courthouse. The courts and other parts of the justice system can operate more productively and less expensively, making court information and records available, and receiving filings, fees, documents, and information, all electronically. These are some of the more obvious and beginning possibilities.

However, these very possibilities also create the risk of worsening old barriers or erecting new barriers to access and causing greater disparities. While the opportunities described seem positive, these innovations assume access to a computer, reasonable proficiency at using the machines, the necessary software programs, reading capability, fluency in English, and sufficient phone or cable and electricity availability. Without all of that, those with the means available get further ahead and those without fall further behind in having the justice system work for them. The lack of equality gets greater, not less.

As a further example, it has been proposed, and increasingly implemented, that some of the laws and regulations that govern us should be published or available only electronically—no more paper copies. The same is true with informational brochures about the courts, legal rights, and procedures and such. This saves money for the government, but makes access to essential information much more difficult for some sectors of society, and the content of law that governs all of us is then available to a select few. Barriers like this already exist, and the trend is continuing with respect to what is meant to be and is called “public information;” the records of some federal agencies are now available only electronically.

Consider also a well-intentioned court-based electronic filing system that is available twenty-four hours per day, seven days per week to those
The Evolution of the ATJ-TBoR

who can use the system and can do so outside of usual work hours, or that gives priority response or action to those who can use that system (persons with a computer, an internet connection, and the skill to navigate what may be a complex software program), but not to those who do not have access outside of usual work hours, or those who do traditional in-person or paper filing.

There has already been some delegation to private companies of the conduct and management of electronic court filing, with associated fees. Additional consideration is being given to privatizing other traditionally public justice system functions such as storage, maintenance, and access to court files and records, the functions of which may be rendered commercially viable by the use of electronic digitalization, maintenance, storage, and search procedures. These practices and considerations anticipate that while the courts and some public agencies would not have to pay for these services, members of the public and those who assist or advocate for them would be charged. Without judging the desirability or lack thereof of any privatization, it is apparent that without careful and enforceable standards prerequisite to any privatization, critical parts and functions of the justice system could well become substantially and disparately inaccessible to many members and segments of the public with significant and damaging consequences.

The foregoing are but a few of the issues and problems that come readily to mind. Many others exist, some recognized, others awaiting study, discovery, and solution.

It has been noted that one of the special aspects of American society is that we are more likely than most societies to appeal to the law and the justice system to correct the denial of legitimate claims to life’s essential resources and rights. Because the justice system is our society’s specific tool to enforce and assure delivery of other essential rights, lack of meaningful or equitable access to the justice system has a geometrically negative effect in perpetuating and increasing the full range of social and economic disparities. On the other hand, increasing meaningful access and reducing or eliminating barriers and disparities in the justice system generates a disproportionately positive effect in providing greater opportunities for the poor and vulnerable, decreasing disparities and divides, and moving society toward greater equality. This is especially true when the values of accessibility and equality are designed from the beginning into major innovations.

It must therefore be emphasized that the ATJ-TBoR initiative is designed to address and move toward solutions not only to justice
system problems, but to fundamental issues of social equity. The transformative capabilities of the new information and communication technologies offer a singular opportunity that simply cannot be overlooked or ignored.

It seems clear that if this effort and its successors are successful in developing and delivering access to the justice system, it will not be an insurmountable task to adapt and use such tools to deliver access to health care and to other essential human needs and services. This cannot be overemphasized.

Further, as indicated earlier, the justice system in our society is a fulcrum. It is a central lever to enforce other essential rights—such as to health care, basic education, and legally mandated food and subsistence. Thus, assuring access to justice is a core requirement that impacts and enhances access to and actual delivery of all other basic and essential human needs.

Consequently, at the same time as the ATJ-TBoR project promulgates a set of authoritative fundamental principles to guide technology development and use in the justice system, it advances the project’s more comprehensive goals and aspirations to enhance access to all the indispensable services and opportunities that every human being needs and should have, including health care, education, basic subsistence (food, clothing, shelter), economic opportunity, safety, and justice.

Such access must be meaningful access, relevant access, access in the community, from libraries, from home, from senior centers, after hours from schools, from community centers, from kiosks in branch government offices or department stores, or malls, and wherever else access needs to be. No longer will the justice system exist solely in the courthouse or lawyers’ offices; no longer will the health care system exist solely in the hospital, clinic, or physician’s office. No longer will the ability to obtain food stamps require a trip of two hours or more duration to and from a large building with a long line, certainly not a private and often a demeaning experience. Access will largely be through the use of information and communication technologies, but we must be and are also interested in the use of any other tools that can help provide or enhance meaningful access to essential opportunities and services.

VI. VALUES, PUBLIC POLICY, AND OPEN PROCESS

This initiative is ultimately about values—no less than our society’s fundamental values—and delivering on those values in realistic,
The Evolution of the ATJ-TBoR

conge, practical, daily ways. That is access to justice—legal justice and social justice.

This initiative is also very importantly about public policy. To date, societal responses to the emerging and rapidly developing information/communication technologies and their use have very largely been reactive rather than proactive. This reactive mode has often resulted in extreme responses, lacking moderation and balance. Identification and analysis of current and potential problems and opportunities and such planning as have occurred have largely been piecemeal and without continuity.

The ATJ-TBoR process has consistently worked to be proactive, inclusive, and deliberative. It was conceived to be and has been an ongoing thoughtful, multi-faceted, careful, and unhurried process of information-gathering, deliberation, and planning with an opportunity to hone and balance the result. It has examined where we are and what is likely coming in order to get ahead of the curve of time and events and provide a means to stay ahead. Coming and continuing change has been carefully planned for, and the very energy and momentum of that change constructively channeled and utilized. Our plan is to ride the wave, not be immersed by it.

The method—a proactive rather than reactive engagement in a multi-disciplinary, deliberative, inclusive, consumer-respectful process and balanced and careful approach to the emerging issues, opportunities, and problems brought about by new technologies—has itself been thought of as an example to follow more generally in formulating public policy around the subject of new or drastically changing concepts, issues, discoveries, conditions, opportunities, and problems.

For the legitimacy, credibility, and quality of both the process itself and its products, we early understood that it is essential that the process enables, receives, listens to, and uses information, viewpoints, and suggestions from people and groups representing a broad array of backgrounds, experiences, perspectives, and expertise, never neglecting to include those the system is meant to serve—its consumers and end users. Outreach and inclusiveness are essential, and from the first day, the project has engaged in outreach and inclusion, and that never stops and will be an ongoing and intrinsic part of our process to the last day. Beginning with this broad vision, ATJ-TBoR leadership set out to include in its committee membership and to work with a range of people, organizations, and efforts that are dealing with technology’s impact on vulnerable populations as well as society in general. From the beginning,
the ATJ-TBoR initiative closely partnered with libraries and librarians, Native American organizations, members and representatives of communities of persons with disabilities, representatives of community technology centers, seniors, organizations working to bring basic telephone and other communication capabilities to all, a range of social service agencies, and members and representatives of other traditionally underserved populations and communities, as well as government agencies, courts, judges, court administrators and clerks, lawyers, technologists, the private sector, and various universities and academics—and more.

To assure that those involved in the project received authentic and practical information and perspectives, along with many other efforts, the Outreach Committee conducted focus groups and interviews with a number of different underserved and diverse groups, including homeless, welfare recipients, persons institutionalized in the correctional system, recent immigrants, farm workers, and victims of domestic violence—and judges. The knowledge gained from the focus groups, the recent Statewide Legal Needs Study, and other direct information sources significantly informed other project committees in their work, and was central in informing the content of the principles of the ATJ-TBoR itself and its accompanying commentary. That knowledge and information will continue to inform other documents, effectuating mechanisms and processes which will enable the ATJ-TBoR project to meet its essential task of assuring the credibility, quality, relevance, and realistic effectiveness of the process and its products.10

We worked with agencies that serve people such as those who were in the focus groups because we understood that as the project planned and then engaged in the process of converting the Access to Justice Technology Principles into practice, it needed collaborators and allies; indeed all groups working on these issues need each other. These collaborations strengthen the likelihood that the combined insights and influence will actually change for the better the way technology is planned, designed, developed, and deployed not only in the justice

10. The report is posted and available to all on the project website at http://www.atjtechbillofrights.org, which itself is an important tool to communicate and interact with various partners and audiences. The statewide Coordinator of the Low Income Telecommunications (LITE) Project is an active member of both the Outreach Committee and the Implementation Committee. The LITE project shares our understanding that access to communication (including basic telephone service) greatly increases the ability of low-income and vulnerable families to access the services they need to build support systems, which in turn allows them to access stable housing, jobs, childcare, and other social services.
The Evolution of the ATJ-TBoR system but also in other core social institutions. The results for the vulnerable and disadvantaged in our communities—and thus for all of us—can only be positive.

VII. STEPS TOWARD DEVELOPMENT AND ADOPTION OF THE CORE PRODUCT: AUTHORITATIVE PRINCIPLES TO GOVERN THE USE OF TECHNOLOGY IN THE JUSTICE SYSTEM

From the beginning it was clear that the Access to Justice Technology Principles (Principles) must be more than words sitting on a library shelf. Those working in the project knew the Principles would be no more than that if the project did not focus on making them live in concrete daily practice in the lives of the people the justice system is supposed to serve.

We knew we must be reality-oriented. Therefore, it has always been viewed and expressed as essential that the Principles be adopted by an authoritative body so that the principles themselves are authoritative, respected, and followed. The project also decided to develop and provide position papers on important areas of practical significance as well as a document to accompany the Principles in the nature of a societal impact statement which identifies the consequences that adoption, enforcement, and implementation of the Principles will likely generate not only for the justice system, but for the broader society and its systems and organizations, such as libraries, community and senior centers, infrastructure needs and such. This document should be the basis for a coherent and balanced plan that will set priorities and shape such consequences. All of these documents are intended to help stimulate and marshal the societal and political will and resources to accomplish these objectives.

For well over a year, and through more than a dozen drafts, the project invited, received, and in many instances incorporated blunt and direct criticism, commentary, suggestions, and edits, from increasingly larger circles of people both inside and outside the justice system community, ultimately arriving at what is likely the final form of the Principles and accompanying comments.11

This has been a very real process of listening and learning. The December 2, 2003 draft—the current draft—of the potential Principles is

the fifteenth, and is a very different document from the early drafts. Every comment and every criticism of each succeeding draft was carefully considered. Some of the earlier drafts were the subject of considerable criticism, and the resulting changes have significantly improved the product. As a result, the same people who criticized now strongly support the current document.

In recent months the project embarked on a path that it hopes and believes will lead to the adoption of these principles in a Court Rule by the Washington State Supreme Court. In these months the content and intent of the Principles have received significant and unanimous support and endorsements. There has been only one noteworthy change, and that change is not to the content or intent, but to the name of the document and the collective Principles from Access to Justice Technology Bill of Rights to Access to Justice Technology Principles.12

Thus, in the proposed Court Rule and other appropriate places, the legal name will be “Access to Justice Technology Principles.”13 That will be accurate, because that is what they are. They will lose nothing in power or effect, and will be true to our Mission Statement adopted at the beginning of this initiative more than two and a half years ago and repeated here: “To create a body of enforceable fundamental principles to ensure that current and future technology both increases opportunities and eliminates barriers to access to and effective utilization of the justice system, thereby improving the quality of justice for all persons in Washington State.”14

12. In the course of the last three years, without trying to do so, the project accomplished what marketing people call “creating a brand.” That brand is “The Access to Justice Technology Bill of Rights” or “ATJ-TBoR.” People are familiar with it, react to it, seem to know what it means, what it stands for. Many will continue to think of and be inspired by these principles by the name by which they first came to know them. The “brand” will continue, on the ATJ-TBoR web site, in everyday discussions and references, in what it means, and by permeating the justice system with the values it stands for.


14. To date the specific and formal endorsements of the current version of the Principles, whether designated as the Access to Justice Technology Bill of Rights or the Access to Justice Technology Principles, are:

On May 10, 2003, the Board of Governors of the Washington State Bar Association unanimously passed a resolution of endorsement of the ATJ-TBoR.

On May 23, 2003, the Judicial Information System Committee of the Supreme Court unanimously endorsed the ATJ-TBoR.

On May 30, 2003, the Council on Public Legal Education unanimously endorsed the ATJ-TBoR.

On June 6, 2003, the ATJ Board, which had unanimously endorsed the ATJ-TBoR at its February 28, 2003, meeting, unanimously passed a formal resolution of endorsement.
The Evolution of the ATJ-TBoR

The ATJ Board currently intends to propose that the court adopt a General Rule declaring the Principles authoritative and applicable to the justice system and its various components. We hope and expect that the court will do so, because the members of the court, themselves demonstrably dedicated to access to justice, know that these Principles must be authoritative, and must become practical reality, providing concrete results in people’s daily lives.

The Principles are now in a condition to be practical and workable tools for decision-makers in the justice system, from judges to administrators to clerks to technologists to webmasters to the Bar Association and many others. They are also a feasible tool for suggesting, planning, prioritizing, implementing, deciding, and providing products, mechanisms, and processes to transform principles into practice, ideas into reality.

VIII. FROM PRINCIPLES TO PRACTICE: EFFECTUATING AND SUPPORTIVE PRODUCTS AND PROCESSES

While the information-gathering, community-building, and development and drafting of the Principles went on, those in the project also undertook the planning and development of products, mechanisms, and processes that will help effectuate the Principles, make them concrete and real in the daily lives of all people in this state—and, it is hoped, elsewhere.

On July 17, 2003, Attorney General Christine Gregoire signed and sent a letter expressing her official and personal support and endorsement of the ATJ-TBoR.

On September 12, 2003, the Washington State Gender and Justice Commission unanimously endorsed the Principles.

On September 19, 2003, the Public Trust and Confidence Committee of the Board for Judicial Administration voted unanimously to send a letter of support of the Principles to the Board for Judicial Administration. This was done on November 6, 2003.

On September 24, 2003, the Board of the Superior Court Judges Association unanimously endorsed the Principles.

On October 3, 2003, the Washington State Minority and Justice Commission unanimously endorsed the Principles.

On October 10, 2003 the Board of the District and Municipal Court Judges Association endorsed the Principles.

On November 14, 2003, the Board for Judicial Administration unanimously endorsed the Principles.

15. On December 5, 2003, the Board of Governors of the Washington State Bar Association endorsed the proposed general rule and voted to join with the ATJ Board with respect to submitting the potential rule to the State Supreme Court.
A. The ATJ Technology Best Practices Template

To bring the Principles to practice, and ensure that access to justice is a reality when technology is used, an ATJ Technology Best Practices Template is currently being developed by a broad-based ATJ-TBoR work group in conjunction with staff at NPower, a premier non-profit organization which helps non-profits with planning and developing their mission-oriented technology. This Template is a key component for actualizing and effectuating the values and principles of access to quality justice.

There are three primary goals of this effort. The first is to assure that all relevant access to justice considerations are taken into account and addressed with optimum quality whenever technology is being planned, designed, developed, or implemented for the justice system. The second goal is to ensure that best or preferred practices in these areas are identified. The third is that information and resources to help incorporate and deliver on these practices are made known and available to the working justice system agency. The template will be comprehensive, and will not only raise questions and issues; it will suggest to the user how best to answer the questions or address the issues, and then identify and link the user to helpful resources to help accomplish the task in the most accessible and economical way feasible. In this way, access will be designed and built into the technology and the system from the start, and as budget people well know, designing and building features into a system from the beginning is far less expensive than having to retrofit them later. This important tool will benefit everyone, from technology designers and programmers to clerks, administrators, judges, lawyers, and, most importantly, the people who need to access and use the justice system.

In mid-2003, the ATJ-TBoR project was awarded a grant by the State Justice Institute to develop and test this Template as applied to electronic filing first to ascertain its capability and effectiveness, and, if capable and effective, then to assure its adaptability to other court systems throughout the nation. Pursuant to the Grant, development and use of the template will then be expanded to apply to other technology tools including document assembly, websites, and public access information terminals. When the Grant ends, and ordinary daily business takes over, other technology tools, as needed, will be added. The Template and its design and content is expected and built to evolve as technology evolves, knowledge and experience in these areas increase, and societal
The Evolution of the ATJ-TBoR

conditions change. The Template will not remain a special project; it will be integrated as a full, regular, and ongoing part of the system.

Although the Template is now focused and designed for use in the justice system, it is also likely to be adapted into a tool to assure access to any service or opportunity. We believe this Template will be readily transferable to other societal systems that provide essential opportunities, services, and products.

B. Sharing Information About Specific Technologies

As those of us engaged in the process have learned about and demonstrated assistive and other technologies that can help persons with disabilities, we have also found that many technologies that can assist those we traditionally think of as persons with disabilities—persons who are sight impaired, hearing impaired, dyslectic, or have difficulty with mobility or coordination—often can also help others we do not usually think of as having a disability. Such people include those with limited literacy, those who use languages other than English, are of different cultures, come from oral traditions, have limited education, or have difficulty with attention or concentration, or just plain lack confidence and are intimidated by the technologies.

Streaming technologies are one application that helps persons with traditional and non-traditional disabilities. Other such technologies we have learned about include satellites, which provide access to underserved and widely disbursed people in rural areas; kiosks, which provide talking information and services in English, Navajo, and Hopi to people with an almost entirely oral tradition of communication; voice internet portals and audible screen readers developed for the blind but helpful for many others who are print impaired although not vision impaired; and interactive kiosks in the community as well as in courthouses which, among other things, enable the filling out and filing of legal forms in a number of areas such as eviction, domestic violence, and the federal earned income tax credit. These kiosks have easy access audio and video instructions, which supplement or supplant print, in English, Spanish, and Vietnamese.

This raised the important issue of the need to provide fully available, accessible, and affordable broadband and high-speed Internet access for reasons other than the commercial and recreational uses usually discussed. Broadband/high-speed Internet access is necessary for the reasonable delivery and use of the technologies that can provide the audio and video graphics and other content and assistive technologies
designed to provide meaningful access to persons with disabilities and other limitations. Other people who, because of limited finances or other reasons, have no computer where they live often use a local library or community center for Internet access when necessary for their important needs. In those circumstances, because of limited numbers of computers and time and staff constraints, to treat all who need access equitably, users are allowed a limited amount of time at the computer. Whatever the time limitation, if any significant amount of material or dense content needs to be transmitted or transacted, a low-speed modem will likely not get the job fully done in the time available. This can be paralyzing and detrimental in the context of legal problems, government services, health care, or job or economic opportunity. High-speed access becomes far more than luxury or fun; it is necessary for equal or at least sufficient opportunity to access essential services or information. The project will express its views and recommendations in greater detail in a position paper on this important issue.

C. Working to Resolve Access and Privacy Issues

The Access and Privacy Work Group will continue its already substantial efforts to help find a proper balance of two constitutional principles that are now competing: the right to open and accessible courts and court records, and the right to personal privacy. But to understand the problem accurately, it must be recognized that there is more here than the issue of access to court information. Critical and largely unconsidered and unaddressed are the major unavoidable issues of what can be done and is already being done with the information in court records that is available to be accessed.

The transforming of court information and individual case records from paper to electronic form and the ability to access such information electronically and remotely cannot be seen as anything less than a transformative and qualitative change in what “access” means to the judicial system and to all those who interact with that system. Whether from court records or from other sources, when enormous amounts of information can be accessed from one’s home or office, or anywhere there is a computer, power, and connectivity, when search capability enables an entire life history to be aggregated in seconds or minutes, when the collected or aggregated information can be disseminated in seconds or minutes to vast numbers of people and institutions (in the hundreds of thousands or millions), or published and put in the public domain permanently for anyone to access at any time, all of this with one
or a few keystrokes, then that is a revolutionary transformation. The consequences for personal privacy boggle the best of minds and cannot be ignored.

What must be done is what courts have historically been called upon to do, and are ultimately good at doing when they address the problem squarely. There must be a thorough and careful effort to find a proper balance between competing fundamental values and rights, or a way to minimize such competition, or both. The courts must do so in this area, whether in their administrative role or in their more traditional role of case-by-case decision-making.

In this context, it is very important to remember the underlying reason for the right to accessible courts and their records. The purpose of the right of access to court information and records has been authoritatively described as to “allow the citizenry to monitor the functioning of our courts, thereby ensuring quality, honesty, and respect for our legal system.”16 It is not for commercial purposes, gossip, or other reasons not relevant to public oversight of the judicial system.

These issues are too central and critical, with too much at stake for too many innocent people (not just litigants, but witnesses, victims, jurors, family members, and others), and for the credibility of the judicial system itself, for this to be handled routinely or “business as usual,” without a careful examination and discussion among thoughtful, deliberative people, to include the many important stakeholders in the system and those who are otherwise involved.

The ATJ-TBoR Committee believes that unless there is some appropriate but not overdone protection of personal privacy and intimate personal, health, financial, and other appropriate information, people will be increasingly reluctant to exercise their right of access to justice, and will avoid the justice system when possible. They will be slow to report crimes, step forward as witnesses, serve as jurors, or use court processes to resolve disputes and legal issues. A disproportionate lack of privacy will chill the exercise of a person’s right to access and use the justice system. These issues are too central, critical, and “new,” with too much at stake for many people, as well as too much at stake for the credibility of the judicial system itself, to be handled routinely or “business as usual.” There must be careful in-depth examination and discussion among thoughtful, deliberative, people, which should include many stakeholders and potentially affected persons and groups. A number of

jurisdictions have taken the approach of a broad-based special task force to study, deliberate, and report on this difficult subject. The ATJ-TBoR project supports such an approach in this state. ATJ-TBoR does not want a closed system; indeed it stands for the opposite. ATJ-TBoR stands for a system that is transparent, open, and accessible, but with careful, balanced privacy and other safeguards. We will continue our efforts along these lines, as will the ATJ Board itself.

D. Collaborating with Libraries and Others to Provide Access

Libraries and librarians have been core partners from the very inception of the ATJ-TBoR project. Jean Holcomb, whose idea and then commitment inspired this project, is the Director of the King County Law Library. Working committee members include representatives of the King County Library System, the Seattle Public Library, the Washington State Library, and the King County Law Library. Libraries have shared with us recent surveys and client feedback, which adds and gives meaning to information gathered from other sources, including our focus groups and the data and analysis from the recent Statewide Legal Needs Survey. Jean and this writer were featured presenters at the 2002 Annual Joint Meeting and Conference of the Oregon/Washington Library Associations, and at the 2003 Washington State Meeting and Conference. Project members also listened and learned at those conferences. We will continue to work closely with librarians and libraries as we implement the principles and make justice available outside the courthouse and in community facilities.

In a demonstration of real-life application of access principles, the project initiated and then worked collaboratively with the State, King County and Seattle Public Library Systems, the King County Law Library, and a number of legal services and private and government social services agencies (including employment offices) to make available in libraries an Internal Revenue Service/Legal Services Corporation sponsored web-based system that enables eligible tax filers to easily fill out the necessary forms for the previously underused Federal Earned Income Tax Credit. The credit can return as much as $4000 a year to a low-income working family.

Currently offered in English, Spanish, and Vietnamese, the tax module asks written questions designed at a fifth-grade literacy level, but also has an audio/video guide person reading the questions aloud in the chosen language. The client answers the questions using a keyboard or mouse at his or her own pace. When all the questions are answered,
The Evolution of the ATJ-TBoR

completed tax forms are available for printing and/or electronic filing. The money comes to the family quickly, thus avoiding the victimization suffered by many families in the past as a result of high priced “Refund Anticipation Loans,” with interest rate ranges from 100% to 700%, offered by many commercial tax preparers.17 This was conceived and initiated by the special people, consistently imaginative and innovative and yet practical, from the Legal Aid Society of Orange County, California, who became close and consistent partners with the ATJ-TBoR project from very early on.

E. Conducting Focus Groups to Prepare for Implementation

The ATJ-TBoR Implementation Committee has embarked on a different series of focus groups stimulated by our having developed the content of the ATJ Technology Principles. This set of focus groups asks questions of many who will be planning, developing, implementing, and instructing about technology use in the justice system, including public and law librarians, law school faculties, court clerks and administrators, and legal services providers to the poor and vulnerable. Faced with the reality of such an authoritative set of principles, the participants are asked what challenges they anticipate facing; what help this may provide them; what opportunities they see and how to realize them; what problems will be presented and how they might deal with them; and what adjustments will be necessary in procedures, training, personnel, funding amounts, and allocations. The responses to these questions are the basis for an in-depth discussion which follows. This process will inform and enable a transition from principles to practice, stimulate logical and realistic planning and prioritization, and hopefully avoid or minimize the institutional fear and resistance which often accompany new standards or requirements. The earlier Outreach Committee focus groups informed the development and creation of the actual content of the ATJ Technology Principles. These current focus groups begin with the content of the Principles and inform their implementation and effectuation.

F. Developing Position Papers on Key Issues

The Implementation Committee is supervising the development of a series of position papers examining and commenting on certain key areas that will influence the translation of values and principles into reality. These include: balancing access to court records and personal privacy; information literacy and usability; high-speed access/broadband; criminal justice issues; and technology use in administrative hearings.

G. Developing a Societal Impact and Optimization Statement

The Implementation Committee is also supervising the development of a comprehensive Societal Impact Statement. This document, theoretically modeled after the more well-known Environmental Impact Statement, will address not only potential negatives, but potential positive opportunities as well. The document will attempt to identify and examine potential consequences of authoritative ATJ Technology Principles, not only in the justice system but also in society in general. For example, the document will consider the effect of the Principles on libraries, community centers, and infrastructure needs. The document has already been outlined and is in process of being drafted, after which it will be circulated, thoroughly critiqued, finalized, and published.

H. Surveying Trial Courts

From its beginning the project participated with the Tribal Technologies group, and has also worked closely with other northwest Native American tribal organizations. Project members have been involved in many tribal meetings and conferences, including the 2002 Tribal Technology Visioning Conference. The ATJ-TBoR Judiciary and Court Administration Committee conducted one of the first ever technology surveys of any state’s tribal courts and court administrations. This survey, designed in part and facilitated by two Native Americans who work in tribal courts and who are also members of the ATJ-TBoR Committee, and administered with the endorsement of the Northwest Tribal Judges Association, discloses the present state of technology in the tribal courts in Washington State, their plans for the next two years, and what they hope for in the future. This will be an important base of information from which to improve both access to tribal justice and the quality of that justice.
The Evolution of the ATJ-TBoR

The Judiciary and Court Administration Committee conducted similar surveys for similar purposes of all the state’s trial courts at all levels. These surveys were of the state’s superior courts; superior court clerks and court administrators; courts of limited jurisdiction, including the district courts and municipal courts; and, as stated above, the Native American tribal courts and their clerks and administrators. The Committee also surveyed private arbitration and mediation agencies and providers.

That Committee is nearing finalization of its report based on the results of the completed surveys it received. The report will present the survey results along with a coherent quantitative and qualitative analysis. This means there will soon be a report and analysis of the state of technology in all levels of the state’s trial courts and the plans and hopes for the future of those who work in those courts. We expect this will provide a realistic basis for the courts to render coherent decisions about improving their technology and its use, as well as for all of us to plan and develop documents, mechanisms, and processes to effectuate and make concrete the ATJ Technology Principles in this core component of the justice system.

I. Conceptualizing an “Ideal” Technology Supported Justice System

The Opportunities, Barriers, and Technology Committee has completed a design of its version of an “ideal” justice system that fully uses technology to create or optimize opportunities and avoid or break down barriers. The design proposes specific ways to use technology to modify current justice system realities so as to effectuate and make real the values and principles of the ATJ-TBoR. It is an aspirational but reality-based document that can be used in ongoing justice system innovation and design, and is published in this issue of the Washington Law Review.

J. Developing a Public Information and Communication Plan

The Steering Committee and the Outreach Committee are cooperating with the ATJ Board’s Communications Committee, the Council for Public Legal Education, and the Public Trust and Confidence Committee of the Board for Judicial Information to develop a public information, education, communication, and marketing plan for communicating information about the ATJ-TBoR to the public. In addition to spreading information about the project, its Principles, and effectuating
mechanisms and processes, this plan will also alert the public to both technological and other available methods of meaningful access to the justice system.\textsuperscript{18}

\textbf{K. Conceptualizing, Encouraging, and Helping Plan for Community Justice Centers}

The project is working on the beginnings of a collaborative process for thinking about and creating community justice centers, which may be independent or with or within other agencies such as law libraries, senior centers, public libraries, or community or ethnic centers. No longer should the justice system exist exclusively in a centralized courthouse. It is clear that technology provides and will increasingly provide substantial opportunities and means to bring significant elements of the justice system out into the communities of those who would or must use it. The process of causing this to happen should soon begin. Like the online Access to Justice Best Practices Template, this too is a major component for actualizing and effectuating the values and principles of access to quality justice.

\textbf{IX. RECOMMENDATIONS FOR THE FUTURE}

This special enterprise known as the Access to Justice Technology Bill of Rights initiative will formally end in January 2004. However, it is clear that this will not just be a special project, the results of which will terminate with its end. Those involved with the project have already begun to accomplish the essential task of institutionalizing what has been done and is to be ongoing, thus assuring that the values the project stands for will permeate our justice system and our society. The Steering Committee will continue to exist after the formal end date. It will oversee the completion of the tasks and projects not yet completed and will also continue to consider and oversee the adoption, institutionalization, and perpetuation of what has been produced and accomplished. This includes identifying those administrative locations where products will be placed or where processes will go, and where oversight responsibility will reside.

One of the accepted recommendations is to institutionalize the structures and habits of thinking whereby from the very beginning

attention is paid and careful consideration given to the incorporation of access to justice concerns in the design, development, and use of court rules, codes, procedures, and practices.

The collaborative work completed over the last year in working with the Judicial Information System Committee and interested stakeholders and parties in the development of Supreme Court General Rule 30 on electronic filing (which, it has been volunteered by persons from other states, is the best they have seen so far) clearly demonstrates this is highly doable. This spirit of consultation, collaboration, and caring about access should become a habit and an institutionalized pattern. There should, for example, be full membership rather than liaison status in the Judicial Information System Committee for an ATJ Board nominee.

Another recommendation is the institutionalization of regular review of the justice system’s operational and procedural rules, codes, and laws—including codes of evidence, confidentiality laws, and codes of professional responsibility and ethics—so as to determine those places where their confluence with technology turns out not to be confluence at all, but collision. Perhaps one or more of the state’s law schools or other academic or research institutions, a designated committee of the Washington State Bar Association, or the Statute Law Committee will take the lead and periodically conduct such a review of the rules, codes, and laws by which the justice system operates.

This effort would seek to identify those places where consideration should be given to making proactive, careful, and balanced adjustments or modifications either of the rule, code, law, or technology, or their application or use so as to avoid potential problems that would discourage user access or damage the usability or quality of accessible justice. Such a situation may occur when, for example, technology use to access confidential legal material is assisted by a library aide, thereby potentially compromising the attorney-client privilege. This effort would also make recommendations to the appropriate person, body, or agency, including courts, legislatures, bar associations, citizen groups, technologists, and others who may have the responsibility or authority to make such adjustments.

X. THE CONFERENCE/SYMPOSIUM AS A CLOSING AND LAUNCHING EVENT

The centerpiece and culmination of the Jurisprudence Committee’s work on the ATJ-TBoR project was a major national symposium and
conference held on January 16 and 17, 2004, entitled, as is this article, “Technology, Values, and the Justice System,” and this dedicated issue of the Washington Law Review.

The event was co-hosted by the Washington State Access to Justice Technology Bill of Rights Committee, the University of Washington School of Law, and the Washington Law Review, and co-sponsored by the University of Washington Information School and the Shidler Center for Law, Commerce, and Technology. The cooperative attitude and action among all of these organizations has been exemplary.

The event included such speakers and presenters as Vinton Cerf, in truth one of the parents of the internet; technology ethicist Helen Nissenbaum of NYU and Princeton; legal historian and professor Mort Horwitz of Harvard Law School; expert in value design in technology Batya Friedman of the University of Washington Information School; Chief Justice Gerry Alexander of the Washington State Supreme Court; Former President of the Legal Services Corporation John McKay; former Chief Justice Robert Utter, currently an adviser on the drafting of constitutions and laws of emerging nations; and many others of substantial expertise and professional stature from the academic to the highly practical and concrete. This will have been the first major event at what was then the recently completed and technologically state-of-the-art William H. Gates Hall, the new home of the University of Washington Law School.

This full edition of the Washington Law Review is dedicated to the proceedings of the conference; the issues addressed and the legal and multidisciplinary issues raised by the ATJ-TBoR; the values, goals, process, and products underlying the ATJ-TBoR; and its consequences and local, national, and international ramifications. The conference and this Washington Law Review symposium edition are not intended to be a conclusion, but rather, as graduations are now called, a commencement. They were designed to stimulate a thoughtful national discussion about these crucial issues, which will become even more important as time goes on.

XI. CONCLUSION

The intention is that the ATJ Technology Principles and its accompanying documents and effectuating products, processes, and mechanisms to bring the principles to reality will result in thoughtful planning, balanced priority setting and decisions, and more efficient and cost-effective solutions. It must also help mobilize the will and resources
to add the powerful tool of information/communication technology to the powerful tool of the law to at last deliver on this nation’s three centuries old promise of accessible, equal and quality justice for everyone.

These Principles are also intended to be a model for other geographical areas and for other essential human service sectors. They are intended and designed to raise the consciousness of the justice system and the broader society to the consideration of access to justice and the use of technology to break down barriers and increase access, thereby minimizing or eliminating both the “Digital Divide” and the persisting centuries-old divides that preceded it.

To accomplish that, we must be clear about how technology is changing the ways in which society conducts its affairs, what the barriers may be, and what solutions are emerging. Then, to bring about and perpetuate the desired changes, we must continue to work in partnership, collaboration, and cooperation with many others. As we build and work with a network of constituencies to enable equal and meaningful opportunity and service delivery, we must also build the collective political will, power, and momentum to assure that the technologies are injected and infused with the values of our federal and state constitutions and other fundamental documents, that they actively serve the professed values of inclusiveness, equality, and justice for all—not only in principle but in fact, not only theoretically but in concrete practical reality. To reach these goals requires commitment, but more is necessary. That commitment must ignite the essential and ongoing requirements of persistence, perseverance, and endurance.

This effort is still the first and to the present only such undertaking anywhere in the nation. As a result, people and organizations from many other places and regions have expressed and acted on their desire to be involved with us and help us produce as good an outcome as possible. They have said they want to help us develop a model that other states, regions, indeed countries, may adapt and use, and they have in fact helped to do that. Our primary job is to create an excellent result for the people of the State of Washington, but if our process and our product is of sufficient quality to help others beyond our borders in the United States or elsewhere, then so much the better. Indeed, the happy fact is that such openness also pays the dividend of enabling us to receive the benefits of the relevant experiences and best ideas of many others, which will in turn benefit the people of our own state. Simply stated, value can be added for everyone; everyone can win.
A great deal has been said and written about what has come to be called “The Digital Divide,” both domestically and internationally. We declare that access to, use of, and respect for the rule of law is an essential way to move to a less divided, more equitable society and world. Accessible, equitable, and quality justice for all individuals and groups is a recognized worldwide value. Access to justice is culture-neutral. Whatever their background and culture, whatever the shape and nature of their system of justice, meaningful access to justice can and does empower people to be part of creating their own just societies. It has been and is the commitment of the Access to Justice Technology Bill of Rights Committee that its work will be a positive contribution to that never-ending effort.
CONCEPTUALIZING THE RIGHT OF ACCESS TO TECHNOLOGY*

Morton J. Horwitz†

I begin with the assumption that my assigned task today is not to defend or justify a wider distribution of access to technology. I accept that as a given. For if that were not true, long years of professorial training in law would lead me, in a Pavlovian manner, to ponder the limits of wider access to technology—its possible negative effects on children or its easy access to the gullible and the innocent, who might become easy prey to the financial and sexual predators among us.

Fortunately, my task is more confined. Accepting that a wider distribution of access to technology is, like wider access to education in general, a social good that is usually to be applauded and promoted, my role is not to defend a broader access to technology but rather to suggest the ways in which an advocate might invoke legal categories and concepts in order to advance that goal. By focusing on legal categories, I should emphasize, I wish to slide past any general moral argument about the injustice of the overall distribution of wealth and how a more just distribution could most efficaciously solve many special problems of unequal access to technology. I am sure that that is true, but you did not have me fly out from Boston to tell you that. Rather, I hope to take on the more limited task of asking whether there are ways of framing the question of access to technology in primarily legal terms.

Beginning our legal inquiry at the most general level, we need first to underline the important ideological difference between the governing assumptions of the U.S. constitutional system and of the constitutional regimes established in most Western European countries after World War II. These post-War constitutions, often called second-generation constitutions in contrast to the eighteenth century American Constitution, express a commitment to the notion of a positive state, a state with affirmative obligations to promote welfare and full

* Remarks made at Technology, Values, and the Justice System, a symposium held on January 16–17, 2004 at the University of Washington School of Law.
† Charles Warren Professor of American Legal History, Harvard Law School.
employment. These constitutions derived either from European social
democratic thought or from the positive liberalism of turn of the century
British and American philosophers like T.H. Green and John Dewey.

By contrast, the prevailing view of the U.S. Constitution is that its
pronouncements are largely negative and meant only to prevent
government from overstepping its bounds, not positive, thereby creating
constitutional duties in the government to maximize the welfare of the
citizenry. In the standard formulation, the American Constitution
embodies eighteenth century ideas of freedom from external restraint;
the twentieth century Western European constitutions, by contrast,
express the idea that real human freedom derives from the maximization
of the talents and abilities of the citizenry and that the state is
constitutionally obliged to promote social conditions that favor the
expansion of positive liberty.

This contrast between “freedom from” and “freedom to” is replicated
in many more technical areas of the law. In that esoteric corner of tort
law dealing with “no duty,” one is struck by the difference between
Europe and America over the duty to rescue. Let us look at the generic
example of the baby lying face down in a puddle, as a stranger cynically
walks by, though he could easily have saved the infant. One may be
surprised to learn that in the United States there is no duty to rescue the
baby, while in most European countries such a duty does exist,
sometimes extending even to criminal liability.

My point is that the distinction between positive and negative duties
extends all the way from top to bottom, from constitutional to tort law,
as a fundamental expression of the difference between European legal
culture and ours. The American attitude is derived from a culture of
rugged individualism and an antipathy to the state. It often rejects ideals

2. See JAMES T. KLOPPENBERG, UNCERTAIN VICTORY: SOCIAL DEMOCRACY AND
PROGRESSIVISM IN EUROPEAN AND AMERICAN THOUGHT, 1870–1920, at 397 (1986); T.H. Green,
ed., 1893); JOHN DEWEY, THE PUBLIC AND ITS PROBLEMS (1927).
REV. 311, 320.
4. Id.
Conceptualizing the Right of Access

of social solidarity, especially when they can only be achieved through enforcement or facilitation by the state as an agent of society.

I began at this level of abstraction in order to highlight the long standing cultural and constitutional obstacles to finding affirmative duties in the law. That does not mean that these obstacles cannot be overcome, but only that legal advocates who strive to craft a right of access to technology need to realize that they are swimming upstream against a vigorous counter current.

At a less abstract, more manageable level of legal discourse, I see four areas of the law that offer promising analogies: the right to education, the right to language, the right to tools, and the right to property. Let me take them up, one at a time. Each of these four rights offers the possibility of analogizing a right of access to technology to a fundamental or near fundamental right already recognized by our legal system.

I. THE RIGHT TO EDUCATION

The broadest and most commonsensical claim would treat a denial of access to technology as equivalent to placing an undue burden on the right to education itself. The strongest version would assert that just as the state has a duty to establish public schools, so too it has a duty to guarantee certain minimum conditions—from adequate books and other learning materials to a physically safe and healthy environment—necessary to keep the basic right to education from being undermined in the first place.

But the current state of the law relating to a constitutional right to education is more complex. Any claim to a strong federal constitutional right to education was defeated in the United States Supreme Court case of San Antonio Independent School District v. Rodriguez. In that case, the Court upheld a system of public school financing based on widely unequal local property taxes against an Equal Protection attack. But even as it was denying the strong claim, the Court did allude to the possibility that an absolute minimum of educational benefits might be constitutionally required. It conceded arguendo that “some identifiable quantum of education is a constitutionally protected prerequisite” to the

8. Id. at 54–55.
9. Id. at 36–37.
“meaningful exercise” of the duties of citizenship, and it acknowledged that the plaintiffs might have a valid claim “if a State’s financing system occasioned an absolute denial of educational opportunities to any of its children.”10 The Court suggested that the case might have come out differently if a “charge could fairly be made that the system fails to provide each child with an opportunity to acquire the basic minimal skills necessary for the enjoyment of the rights of speech and of full participation in the political process.”11

Despite these concessions, after Rodriguez, the claim that unequal funding between poor and rich school districts violates Equal Protection shifted to the states under their own equal protection clauses. However, one other subsequent Supreme Court decision did find that it was an Equal Protection violation for Texas to refuse public education to children of illegal aliens.12 Though constrained by the Rodriguez decision holding that public education was not a fundamental constitutional right, the Court formulated what Professor Tribe calls “a hybrid equal protection test” that highlighted, not a “fundamental right” to education but rather the “fundamental role” education plays “in maintaining the fabric of our society.”13

What scholars have called the “second wave” of school finance reform litigation, which involved claims brought under state constitutions, began after Rodriguez was decided in 1973. In the same year, the New Jersey Supreme Court initiated the second wave by invoking provisions of the New Jersey constitution to strike down the state’s system of educational funding.14 The second wave of litigation had some success but more frequent failure, and as time passed, success became a rarity as failures predominated.15

The waning success of state equal protection decisions recognizing plaintiffs’ fundamental right to equal education gave way to a “third wave” during the 1990s that focused not on inequality but on whether a

10. Id.
11. Id. at 37.
Conceptualizing the Right of Access

child had received “an adequate education.” 16 “Although often coupled with equal opportunity claims, the third wave’s hallmark is challenging the adequacy of education rather than the equality of financing.” 17 Plaintiffs turned away from state equality clauses and framed their claims under other state constitutional provisions. One group of state constitutions specifies explicit quality standards for public education. The Illinois constitution, for example, declares that “[a] fundamental goal of the People of the State is the educational development of all persons to the limits of their capacities.” 18 It requires the state to “provide for an efficient system of high quality public educational institutions and services.” 19

Montana requires an educational system that “will develop the full educational potential of each person.” 20 The Virginia constitution requires the legislature to “ensure that an educational program of high quality is established and continually maintained.” 21 Louisiana must provide learning experiences that are “humane, just and designed to promote excellence in order that every individual may be afforded an equal opportunity to develop to his full potential.” 22

The Washington constitution requires that the state make “ample provision” for the education of all children. 23 In one of the first “third wave” standards cases, the Washington State Supreme Court in 1978 held that the clause was not a mere suggestion but a specific duty imposed on the legislature. 24 After defining the constitutional word “ample” as “liberal, unrestrained, without parsimony, fully, [and] sufficient,” 25 the court found that the state’s constitutional duty “embraces broad educational opportunities needed in the contemporary

18. ILL. CONST. art. 10, § 1.
19. Id.
20. MONT. CONST. art. X, § 1.
21. VA. CONST. art. VIII, § 1.
22. LA. CONST. art. VIII pmbl.
23. WASH. CONST. art. IX, § 1.
25. Id. at 516, 585 P.2d at 93.
setting to equip our children for their role as citizens and as potential competitors in today’s market as well as in the marketplace of ideas.”

There are, of course, many other state constitutional provisions that are either weaker or vaguer than the ones just considered. In one of these states, North Carolina, there is a remarkable decision under a relatively weak constitutional provision. The North Carolina constitution requires that the legislature support “a general and uniform system of free public schools . . . wherein equal opportunities shall be provided for all students.” In Leandro v. State, the North Carolina Supreme Court held that all children are entitled to the same minimum qualitative level of education, regardless of which schools the children attend. The court found that the North Carolina constitution guarantees a “sound basic education.” The court also found that “an education that does not serve the purpose of preparing students to participate and compete in the society in which they live and work is devoid of substance and is constitutionally inadequate.” A “sound basic education,” the court said, consists of the opportunity to develop the following four skills: (1) sufficient ability to read, write, and speak English, and a fundamental knowledge of mathematics and physical science; (2) a fundamental knowledge of geography, history, and basic economic and political systems; (3) sufficient skills to enable students to engage successfully in further education or vocational training; and (4) sufficient skills to allow students to compete equally with others in further education or employment.

The North Carolina legislature has actually passed a statute that defines a “sound basic education” to include “the areas of the arts, communication skills, physical education and personal health and safety, mathematics, media and computer skills, science, second languages, social studies, and vocational and technical education.”

This third wave shift to adequacy standards has brought into view a large variety of state constitutional provisions that are much better suited to the task of establishing a right of access to technology than the more

26. Id. at 517, 585 P.2d at 94.
27. N.C. CONST. art. IX, § 2.
29. Id. at 254.
30. Id.
31. Id.
32. Id. at 255.
33. See Packard, supra note 17, at 1507–08 (emphasis added).
Conceptualizing the Right of Access

potentially sweeping equality provisions. Moreover, because, unlike the federal government, states have plenary power over education, there are many state constitutional and statutory provisions recognizing the sort of broad affirmative duties in the state that are as sweeping as those I identified at the outset with post-World War II second-generation European constitutions.

II. THE RIGHT TO LANGUAGE

For many courts, however, we are still in need of a narrower analogy than the right to education provides. Perhaps the analogy between access to technology and language rights provides a still tighter fit. The United States Supreme Court has held that children whose mother tongue is not English have a right to learn English, and the public schools have a duty to provide the opportunity to achieve literacy in English.\(^\text{34}\) We have just seen that many states have established an affirmative duty to attain literacy in English.\(^\text{35}\)

It is not difficult to conceive of access to information technology as access to a new kind of primary language that, like English, is an indispensable prerequisite for becoming a full participant in modern society and economy. In order for one to be actually in a position to learn and grow in twenty-first century America, access to the language of technology seems no less fundamental than access to knowledge of the English language itself.

The analogy between language and information technology has the great advantage of narrowing the right to education argument and of steering it away from the always potentially lethal slippery slope of an affirmative duty of equality. Instead, the analogy focuses on access to a more limited class of language skills that can be said to provide a unique window into participation in the culture. There might be arguments about whether there are other competitors such as mathematics, the sciences, or economics that actually do perform as fundamental a gatekeeping role as the primary language of the culture. My assumption is that the strongest claimant, after English itself, to the status of a primary language that controls access to the world is the language of information technology.

\(^{35}\) See supra notes 18–33 and accompanying text.
What do I mean when I speak of access to information technology? How do we contain the potentially overbroad meanings that might sink the whole enterprise? For example, a right of access to information technology might range all the way from the claim that the state needs to provide students with their own computers to the narrower claim that every school needs to make children familiar with the workings of a computer. A requirement of computer “literacy” might be an intermediate goal, and it might be tested in the same ways as we measure literacy in ordinary language.

To apply the idea of a right of access to technology in a different setting, let us shift from the school to the criminal justice system. In many parts of the country, criminal defense attorneys have substantially less access to expensive computer research than do state-funded prosecutors. One way to think about the problem might be first to recognize that the constitutional right to counsel does not include a guarantee that defendants must be represented by equally skilled or well-trained lawyers. Except for very wealthy defendants, our criminal justice system regularly ignores the dramatic asymmetry in resources between prosecutors and defendants. Only when the quality of the defendant’s representation falls below the standard of “adequacy” does the law step in to correct the imbalance.

It would not be far fetched to claim that access to computer legal research should be an important element in judging the adequacy of a lawyer’s defense. Would that also be true if the defense lawyer were a computer illiterate who had access to a complete print collection of legal materials? If we have permitted major inequalities between the state and the defense in their access to print materials, is there any reason why we should require more equality in access to computer technology?

The claim that lack of access to computer research is different in kind from many other kinds of inadequate lawyering may be hard for a court to swallow. The existing legal standard encourages judges to look the other way except in the most egregious cases of inadequate representation. To create a kind of per se rule that elevates access to information technology to a preeminent position in measuring adequate lawyering might easily create the sort of slippery slope paranoia that sees access to computer information as the entering wedge in a more radical effort to equalize lawyering.

If I had a better grasp of both the full range and the complexities of current technological possibilities, I might be able to spin out even more mind-boggling choices between broad and narrow versions of the
Conceptualizing the Right of Access

meaning of the right of access to information technology. The important point, however, is to stay close to the basic idea that in our culture, both language and information technology uniquely control and structure our ability to learn about and participate in the world.

III. THE RIGHT TO TOOLS

If either education or language rights fail as analogies either because they are not intuitively compelling or because they threaten to lead to overbroad conclusions, a third set of analogies might prove more compelling. It begins by conceiving of access to information technology as a tool, both literally and metaphorically. In the literal sense, one would want to draw on the underlying policy contained in state statutes that exempt mechanics’ tools from seizure in debt or bankruptcy proceeding. These statutes, adopted in most states after the Revolution, were based on the sensible idea that if workers’ tools could be seized, workers would be deprived of the very means of staying out of insolvency in the future. Our interest in tools as an analogy is that it expresses the idea that the acquisition of certain things like “tools” may be specially protected because they are regarded as a prerequisite to workers participating in a market economy without losing all of their actual autonomy or agency.

To extend this analogy, can books be thought of as tools? Are there similar exemptions for, say, the professional library of a doctor, lawyer, minister, or teacher? From books, it would be a short step to including information technology in the privileged circle of tools.

IV. THE RIGHT TO PROPERTY

I have saved for last the most prominent way in which contemporary legal discourse structures the question of a right of access to information technology.

36. See 3–4 ANNUAL LAW REGISTER OF THE UNITED STATES (William Griffith ed., 1822) (establishing that by the 1820s, two-thirds of the states had already enacted statutory provisions that exempted the tools of one’s trade from execution in bankruptcy).

37. See id. at 503, 664, 681–82, 991 (establishing that by the 1820s, four states exempted books from execution in bankruptcy: Louisiana (“books of professional men”); Maine and Massachusetts (“bibles and school books”); Mississippi (“books of a student”); see also Bernard R. Trujillo, The Wisconsin Exemption Clause Debate of 1846: An Historical Perspective on the Regulation of Debt, 1998 WIS. L. REV. 747, 757 (1998) (noting that the first version, not adopted, of Wisconsin’s Exemption Clause in its 1846 Constitution provided for a $500 exemption from taxation and execution of, inter alia, mechanics tools, farming utensils, and professor’s books).
technology. In a word, the discussion overwhelmingly centers on questions of property rights. A right of access is conceived of as the reciprocal of the right to property. By determining the property rights of companies that sell music on CDs, you also determine the absence of any property rights in those who wish to download that music for free on Napster. By expanding the scope of intellectual property rights, courts have thus reduced access to information technology. A clear example is the “fair use” exception to the claim of infringement in the copyright law. As the scope of property rights expands—in this case, as the fair use exception is narrowed—access to technology is less possible.

My colleague, William Fisher, has written a short history of intellectual property law, by tracing the evolution of its different technical branches—copyright, patents, unfair competition, trademarks, trade secrets, business goodwill. Fisher has concluded that a substantial expansion of property rights has occurred in the recent past, say since World War II, especially as each of these sub-fields, with their own specialized sub-rules, have been blended into a new, more generalized and abstract generic category known as intellectual property. Fisher sees the expansion of intellectual property rights as deriving from both changes in the economy—the shift from an agricultural to industrial to information-based society—as well as significant early ideological commitments to a Lockean labor (just deserts) theory of property. Important cultural factors he identifies as supporting the expansion of intellectual property rights are the Romantic ideal of authorship and the image of the inventor as a creative genius. Finally, he identifies a gradual shift in legal terminology that has recently created the generic field of intellectual property, which has contributed to the “propertization” of the field. For example, in many of these fields expansive language of property rights has displaced the traditional discourse of limited monopoly, which had placed a stronger burden on the property-claiming plaintiff.

It seems to me that, for better or for worse, the immediate battle over access to technology will continue to be framed primarily as a question of property rights. Therefore, it is necessary for advocates of increased

access to be fully attuned to the possibility of expanding access by limiting the scope of property rights. Or, to put the matter more precisely, even when the issue of access is framed as a question of property, there is no reason to concede that property rights are absolute. While the rhetoric of property rights since the 1980s has increasingly taken on an absolutist spin, in actuality, most traditional conceptions of property rights already have incorporated traditional limitations that had been originally created out of a sense of the public interest. A familiar, albeit special, example is the fair use defense in copyright. Against a constitutional background that posited that the grant of a copyright was a limited privilege created in the public interest for the purpose of encouraging the dissemination of useful knowledge, it was relatively easy to conceive of a fair use exception as included within the statutory definition of the copyright grant. In other areas of property law, because the limitations on the use of property are more deeply embedded in the technical law and thus less visible, one finds more unrestrained use of the rhetoric of absolute property rights.

In fact, even the most absolute-sounding subject of trespass to land can be shown to be riddled with exceptions. From the Supreme Court’s reluctance during the Civil Rights Era to enforce the trespass laws against civil rights demonstrators engaged in sit-ins at segregated private facilities to state courts entertaining the defense of necessity put forth by activists protesting against the dangers from nuclear power facilities, we see the malleability of property rights when defenses of necessity or of an implied public easement are used to limit the absolutist sound of trespass.

It is necessary to dwell a moment more on trespass not only because it is often misleadingly deployed to symbolize the absoluteness of property claims but also because it has been increasingly invoked by courts as a central basis for limiting access to information technology. Not only is the hacker found to be a trespasser even without proof of the high degree of intentionality that suing in trespass often requires, but trespass has been deployed to enjoin a former employee from writing emails to his

ex-coworkers critical of his former employer.\textsuperscript{40} Despite important First Amendment issues, a court held that this was a trespass onto the company’s server, which entitled the company to deny the critical ex-employee access to the company’s email.\textsuperscript{41} It appears that whenever a court analyzes a denial of access to computer technology in terms of trespass, by framing the issue as analogous to real property containing more or less clearly defined boundaries, the court is already predisposed to adopt a mental picture of boundary-crossing defendants who are prima facie guilty of trespass onto plaintiff’s fee simple. However, perhaps the greatest difference between land and intellectual property is that the latter does not offer the kind of clear physical boundaries generally present in the traditional trespass case. It is the very issue of how to set the more intangible boundary of the intellectual property right that is really in dispute. So if one is dragged into thinking in these overly literal terms involving trespass, we should at least be aware that technical trespass law also contains public interest limitations on private property rights in the forms of public easements or the defense of necessity. These terms often express presumptions about the importance of carving out of the trespass doctrine exceptions that protect public space or public ways from the total control of private property owners.\textsuperscript{42} If the trespass analogy applies to intellectual property, the public interest limitations should also apply.

The explosion of interest in intellectual property during the past twenty years has also created an intellectual explosion in law schools. Beyond raising difficult questions about how to apply legal categories to rapidly changing technology, it is also producing an intellectual explosion over theories of property that last took place almost one hundred years ago. It has reopened a discussion that the Legal Realists began during the 1920s about the socially constructed character of property rights.\textsuperscript{43} In a word, the Legal Realists persuaded a generation of scholars that property rights are not natural rights but rather socially created privileges established for social purposes.

\textsuperscript{40} Intel Corp. v. Hamidi, 114 Cal. Rptr. 2d 244, 247–52 (Ct. App. 2001), rev’d, 71 P.3d 296 (Cal. 2003).
\textsuperscript{41} Id. at 252, 255.
Conceptualizing the Right of Access

The advocate of expanded access to technology faces a difficult challenge. In an era in which science and technology have demonstrated a fantastic ability to create and to innovate, it is not easy to resist the argument that innovation requires unrestricted incentives in the form of property rights.

But the history of technology also demonstrates the ways in which those who were awarded monopolies through copyright, patent, or trademark sought to extend their advantages beyond the time period and privileges they were granted. The recent profligate extension of copyright protection bears little relationship to creating future incentives.44 Instead, it needs to be seen as a triumph of interest group politics. The pharmaceutical companies use every questionable device to extend their patents and intimidate generic drug companies. The pharmaceutical example illustrates a more general point about the process of granting property rights. It enormously strengthens the market power of first entrants who continue to exercise disproportionate market power even after their monopolies expire.

PRIVACY AS CONTEXTUAL INTEGRITY

Helen Nissenbaum*

Abstract: The practices of public surveillance, which include the monitoring of individuals in public through a variety of media (e.g., video, data, online), are among the least understood and controversial challenges to privacy in an age of information technologies. The fragmentary nature of privacy policy in the United States reflects not only the oppositional pulls of diverse vested interests, but also the ambivalence of unsettled intuitions on mundane phenomena such as shopper cards, closed-circuit television, and biometrics. This Article, which extends earlier work on the problem of privacy in public, explains why some of the prominent theoretical approaches to privacy, which were developed over time to meet traditional privacy challenges, yield unsatisfactory conclusions in the case of public surveillance. It posits a new construct, “contextual integrity,” as an alternative benchmark for privacy, to capture the nature of challenges posed by information technologies. Contextual integrity ties adequate protection for privacy to norms of specific contexts, demanding that information gathering and dissemination be appropriate to that context and obey the governing norms of distribution within it. Building on the idea of “spheres of justice,” developed by political philosopher Michael Walzer, this Article argues that public surveillance violates a right to privacy because it violates contextual integrity; as such, it constitutes injustice and even tyranny.

I. INTRODUCTION

Privacy is one of the most enduring social issues associated with information technologies. It has been a fixture in public discourse through radical transformations of technology from stand-alone computers, housing massive databases of government and other large institutions, to the current distributed network of computers with linked information systems, such as the World Wide Web, networked mobile devices, video and radio-frequency surveillance systems, and computer-enabled biometric identification. Among many privacy controversies that have stirred public concern, a particular set of cases, to which I have applied the label “public surveillance,” remains vexing not only because these cases drive opponents into seemingly irreconcilable stances, but because traditional theoretical insights fail to clarify the sources of their controversial nature.1 This Article seeks to shed light on the problem of

* Associate Professor, Department of Culture & Communication, New York University, East Building 7th Floor, 239 Greene Street, New York, New York 10003. E-mail address: helen.nissenbaum@nyu.edu.

Many people and institutions have inspired and helped me in this endeavor, beginning with the Institute for Advanced Study, School of Social Sciences, where I wrote and presented early drafts.
public surveillance first by explaining why it is fundamentally irreconcilable within the predominant framework that shapes contemporary privacy policy, and second by positing a new concept—contextual integrity—to explain the normative roots of uneasiness over public surveillance. This Article’s central contention is that contextual integrity is the appropriate benchmark of privacy. Before taking up these general points, it is useful first to consider a few specific illustrations of public surveillance.

Case 1: Public Records Online. Local, state, and federal officials question the wisdom of initiatives to place public records online, making them freely available over the Internet and World Wide Web. The availability to citizens of public records, such as arrest records; driving records; birth, death, and marriage records; public school information; property ownership; zoning and community planning records; as well as court records, serves the unquestionable purpose of open government. Nevertheless, the initiatives to move these records online in their entirety, making them even more accessible, cause unease among many, including government officials and advocacy organizations, such as the National Network to End Domestic Violence and the American Civil Liberties Union.

State supreme courts, for example, with jurisdiction over court records, are mindful of concerns raised by advocates of victims of domestic violence and other crimes, among others, who point out the dangers inherent in these new levels of accessibility. Yet their worries seem paradoxical. The records in question are already publicly available. Computerizing and placing them online is merely an administrative

Drafts were further sharpened through opportunities to present at colloquia and workshops held at the New Jersey Bar Association, Princeton University’s Program in Law and Public Affairs, University of British Columbia, University of California, San Diego, University of Maryland, University of Washington, and the Social Science Research Council. Colleagues who have shared essential insights and expertise include Grayson Barber, Rodney Benson, Aaron Goldberg, Jeroen van den Hoven, Natalie Jeremijenko, Bob Salmaggi, Bilge Yesil, and Michael Walzer. I received outstanding research and editorial assistance from Danny Bloch, Rachel Byrne, and Brian Cogan. Grants from the Ford Foundation (Knowledge, Creativity, and Freedom Program) and National Science Foundation (SBR-9729447 and ITR-0331542) have supported my research as well as the writing of this Article.


2. See Robert Gellman, Public Records—Access, Privacy, and Public Policy: A Discussion Paper, 12 GOV’T INFO. Q. 391 (1995) (noting that restrictions do apply on access to government records). The point here is whether any changes are necessary in the transition from paper-based access to online access to these records.
Privacy as Contextual Integrity

move towards greater efficiency. Nothing has changed, fundamentally. Are these worries rational? Is there genuine cause for resistance?

Case 2: Consumer Profiling and Data Mining. Most people in the United States are aware, at some level, that virtually all their commercial activities are digitally recorded and stored. They understand that actions such as buying with credit cards, placing online orders, using frequent shopper cards, visiting and registering at certain websites, and subscribing to magazines leave digital trails that are stored away in large databases somewhere. Fewer are aware that this information is shipped off and aggregated in data warehouses where it is organized, stored, and analyzed. Personal data is the “gold” of a new category of companies, like Axiom, that sell this information, sometimes organized by individual profiles, to a variety of parties, spawning product, subscriptions, credit card, and mortgage offers, as well as annoying phone solicitations, special attention at airport security, and targeted banner and pop-up advertisements. When the popular media writes about these webs of personal information from time to time, many react with indignation. Why? Often the information in question is not confidential or sensitive in nature.

Case 3: Radio Frequency Identification (RFID) Tags. These tiny chips—which can be implanted in or attached to virtually anything from washing machines, sweaters, and milk cartons to livestock and, it is anticipated, one day, people—are able to broadcast information to radio signal scanners up to ten feet away. Although prospective users of these tags have lauded their tremendous promise for streamlining the stocking, warehousing, and delivery of goods, as well as in preventing theft and other losses, privacy advocates point out a worrisome possibility of a multitude of commodities with the capacity to disseminate information about consumers without their permission or even awareness. Why does this worry us? After all, information will be gathered mainly from open or public places where the powerful radio frequency emitters would most likely be located.

All three cases are spurred by technological developments and developments in their applications that radically enhance the ability to collect, analyze, and disseminate information. Case 1 highlights how

---

3. It is important to note that we are not adopting a deterministic model either of technological development or of technology’s impact on society. When we say that a technological development or an application of technology has had particular results, we assume an undeniably complex backdrop of social, political, economic, and institutional factors that give meaning, momentum, and direction to observed outcomes.
great increments in the ability to disseminate and provide access to information prompt disquiet, particularly at the prospect of local access giving way to global broadcast. This worry seems to be a contemporary version of the one evoked in Samuel Warren and Louis Brandeis’ seminal work calling for a right of privacy in the face of then-new developments in photographic and printing technologies.4

In Case 2, it is advances in storage, aggregation, analysis, and extraction (mining) of information both online and off-line that spur questions.5 One of the earliest cases to spur a grass-roots, Internet-mediated storm of protest centered on Lotus Marketplace: Households, a database intended for distribution on CD-ROMs. The database contained aggregated information about roughly 120 million individuals in the United States, including names, addresses, types of dwelling, marital status, gender, age, approximate household income, and so forth. Eventually, the two companies collaborating on the venture, Lotus Development and Equifax Inc., backed off, citing negative publicity.6

Case 3 focuses attention on enhanced modes of gathering or capturing information as in automated road toll systems like EZ Pass, video surveillance and face recognition systems, web browser cookies, biometrics, thermal imaging, and more.7

One could read these cases simply as public policy disputes in which groups with opposing interests face off against one another, each seeking to promote its own goals, desires, preferences, and interests above those of opponents in the dispute.8 This reading is not entirely unproductive as


5. See Laura J. Gurak, Persuasion and Privacy in Cyberspace: The Online Protests over Lotus Marketplace and the Clipper Chip (1997). Another case that has touched off a flurry of concern and protest is profiling of online advertising companies, such as Doubleclick, that monitor the online web-surfing behaviors of millions of users, frequently merging online records with other information about these users. See the website of the Electronic Privacy Information Center for a full account of this case at http://www.epic.org (last visited Jan. 17, 2004).

6. See Nissenbaum, supra note 1.


8. See Priscilla M. Regan, Legislating Privacy: Technology, Social Values, and Public Policy (1995) (providing a rich reading of many interest based privacy disputes during the
Privacy as Contextual Integrity

it at least requires an understanding of how technologies can affect
diverse social groups differentially and how these differences suggest
particular reactive policies, which in turn have the capacity to shape
further technical developments.

In this Article, however, the fluctuations of public interest politics,
public policy, and at times law, are not central; the focus, rather, is the
foundation for policy and law expressed in terms of moral, political, and
social values. We will not be pursuing or presenting specific policies and
strategies for achieving them, but trying to explain, systematically, why
particular policies, laws, and moral prescriptions are correct. Another
way of saying this is that our purpose is to articulate a justificatory
framework for addressing the problem of public surveillance including
the many disputes typified by our Cases 1, 2, and 3 above. Such a
framework would not only address specific cases before us, but would
allow them to serve as precedents for future disputes in a way that Lotus
Marketplace: Households, despite its successful outcome, never did. A
justificatory framework linking cases across time provides rationality to
their resolution that rises above the power plays of protagonists and
antagonists.9

Before proceeding, it is necessary to define boundaries and
terminology. The scope of privacy is wide-ranging—potentially
extending over information, activities, decisions, thoughts, bodies, and
communication. A full theory of privacy would need to take account of
all these dimensions, even if, eventually, it asserted theoretically
grounded exclusions. Such is frequently the case for accounts of privacy
that do not, for example, consider the right to abortion as a component of
a right to privacy.10 The goals of this Article are more limited, not
aiming for a full theory of privacy but only a theoretical account of a

period roughly from 1890 through 1991); see also Susannah Fox, The Pew Internet & Am. Life
Project, Trust and Privacy Online: Why Americans Want To Rewrite the Rules (2000)
(survey of popular privacy preferences), available at
http://www.pewinternet.org/reports/pdfs/PIP_Trust_Privacy_Report.pdf; Joseph Turow,
Annenburg Pub. Policy Ctr. of the Univ. of Pa., Americans & Online Privacy; The

9. This is in contrast with the case of Lotus Marketplace: Households, where privacy advocates
arguably “won” but not in a precedent setting way in the current landscape of data collection,
aggregation, and analysis.

10. This is sometimes called “constitutional privacy.” For discussion of the full picture and
opposing views, see Anita L. Allen, Uneasy Access: Privacy for Women in a Free Society
(1988); Judith Wagner DeCew, In Pursuit of Privacy: Law, Ethics, and the Rise of
Technology (1997); Ruth Gavison, Privacy and the Limits of Law, 89 Yale L.J. 421 (1980).
right to privacy as it applies to information about people. Furthermore, it undertakes this aim in relation to individual, identifiable persons—not taking up questions about the privacy of groups or institutions. Finally, for purposes of precision, we will reserve the term “personal information” for the general sense of information about persons; “sensitive” or “confidential” will indicate the special categories of information for which the term “personal information” is sometimes used.

The balance of this Article is divided into two parts. The first part posits and discusses a framework consisting of three conceptually independent principles that define an approach to privacy protection that dominates contemporary public discussion, policy, and legal landscape.11 It includes subparts devoted to each of the principles, respectively,12 and a subpart on contentious cases in which opposing sides disagree on whether given principles apply to the cases in question.13 The final subpart explains why public surveillance is problematic for this three-principle framework. Unlike the contentious cases discussed before, public surveillance seems to fall entirely outside its range of application.14

The second part of this Article proposes an alternative account of privacy in terms of “contextual integrity”—an introduction to the layer of social analysis upon which the idea of contextual integrity is built.15 Developed by social theorists, it involves a far more complex domain of social spheres (fields, domains, contexts) than the one that typically grounds privacy theories, namely, the dichotomous spheres of public and private. Following this introduction, the first two subparts describe, respectively, two “informational norms” that govern these contexts of social life, namely, appropriateness and distribution.16 The third subpart, anticipating challenges to the normative force of contextual integrity, gives an account of its normative foundations.17 The fourth subpart shows how contextual integrity may be applied to the three Cases described in this Article’s introduction, showing that it easily captures

11. See infra Part II.
12. See infra Part II.A–C.
13. See infra Part II.D.
14. See infra Part II.E.
15. See infra Part III.
16. See infra Part III.A–B.
17. See infra Part III.C.
Privacy as Contextual Integrity

their problematic roots. In the final subpart, the approach to privacy through contextual integrity is contrasted with other theoretical approaches that also extend beyond the three-principle framework.

II. THREE PRINCIPLES

The search for a justificatory framework is a search for theories and principles that yield reasons for favoring one general policy or another and for resolving particular cases. It is useful to understand why prevailing principles that have guided so much of contemporary privacy policy and law in the United States offer little guidance in many hard cases, including the three described at the beginning of this Article. Surveying the fields of public policy development, regulation and statutory law, court decisions, and social and commercial practices during the twentieth century we find that three principles dominate public deliberation surrounding privacy. The three principles are concerned with: (1) limiting surveillance of citizens and use of information about them by agents of government, (2) restricting access to sensitive, personal, or private information, and (3) curtailing intrusions into places deemed private or personal.

A. Principle 1: Protecting Privacy of Individuals Against Intrusive Government Agents

This principle comes into play when questions arise about intrusions by agents of government (or government agencies or representatives) who are accused of acting overzealously in collecting and using personal information. This principle can be understood as a special case of the powerful, more general principle of protecting individuals against unacceptable government domination. Privacy is thus protected by reference to general, well-defined, and generally accepted political principles addressing the balance of power, which, among other things, set limits on government intrusiveness into the lives and liberty of individuals. Data gathering and surveillance are among many forms of government action in relation to individuals needing to be stemmed.

In the United States, the Constitution and Bill of Rights provide what is probably the most significant source of principles defining limits

18. See infra Part III.D.
19. See infra Part III.E.
20. U.S. CONST. amends. 1–X.
to the powers of federal government in relation to the liberty and autonomy of individuals and individual states. They also serve as a powerful reference point for privacy protection. Although, as commonly noted, the U.S. Constitution does not explicitly use the term “privacy,” many legal experts agree that various aspects of privacy are, in fact, defended against government action through several of the amendments, including the First (speech, religion, and association), Third (quartering soldiers), Fourth (search and seizure), Fifth (self-incrimination), Ninth (general liberties), and even the Fourteenth (personal liberty versus state action) Amendments. The U.S. Constitution, as we know, draws on other tracts, including English common law and works of the great political philosophers that have contributed fundamentally to defining the powers and limits of governments in democratic societies embraced not only in the United States, but in the laws and political institutions of western democracies and many beyond.\(^\text{21}\)

Not all legal restraints on governmental gathering and use of information about individuals stem from the Constitution. Others have been expressed in state and federal statutes, with a notable peak of activity in the mid- to late 1960s, coinciding with a steady increase in the creation and use of electronic databases for administrative and statistical purposes.\(^\text{22}\) Priscilla Regan’s detailed account of privacy policy from the 1960s through the 1980s suggests that informational privacy became a topic of intense public scrutiny around the late 1960s following a proposal in 1965 by the Social Science Research Council to create a Federal Data Center to coordinate centrally the use of government


\(^{22}\) For discussions of the trend toward increasing reliance upon computerized record-keeping systems by government and other agencies, see, for example, COLIN J. BENNETT, REGULATING PRIVACY: DATA PROTECTION AND PUBLIC POLICY IN EUROPE AND THE UNITED STATES (1992); DAVID BURNHAM, THE RISE OF THE COMPUTER STATE (1983); DAVID H. FLAHERTY, PRIVACY AND GOVERNMENT DATA BANKS: AN INTERNATIONAL PERSPECTIVE (1979); KENNETH C. LAUDON, DOSSIER SOCIETY: VALUE CHOICES IN THE DESIGN OF NATIONAL INFORMATION SYSTEMS (1986); GARY T. MARX, UNDERCOVER: POLICE SURVEILLANCE IN AMERICA (1988); REGAN, supra note 8; JAMES B. RULE, PRIVATE LIVES AND PUBLIC SURVEILLANCE (1973); Richard P. Kusserow, Fighting Fraud, Waste, and Abuse, 12 BUREAUCRAT 23 (1983); James B. Rule et al., Documentary Identification and Mass Surveillance in the United States, 31 SOC. PROBS. 222 (1983).
Privacy as Contextual Integrity

statistical information. This culminated in the Privacy Act of 1974, which placed significant limits on the uses to which agencies of federal government could put the databases of personal information. Many other statutes followed that placed specific restrictions on government agents in their collection and use of personal information.

For purposes of our discussion, more relevant than the specific details about legal restrictions on government agents is the general source of momentum behind these restrictions, in particular, a principled commitment to limited government powers in the name of individual autonomy and liberty. To the extent that protecting privacy against government intrusion can be portrayed as an insurance policy against the emergence of totalitarianism, the rhetoric of limiting government powers can be parlayed into protection of privacy. During the 1950s until the end of the Cold War, when regimes in the East loomed vividly in public consciousness and fictional constructions, like George Orwell’s Big Brother in 1984, entered the public imagination, the U.S. Department of Health, Education, and Welfare’s Secretary’s Advisory Committee on Automated Personal Data Systems found a receptive audience for their seminal 1973 report on the impacts of computerized record-keeping on

23. See Regan, supra note 8.
25. Id. We should not exaggerate the scope of success. The Privacy Act of 1974 addressed only government record-keeping, bowing to the lobbying of large private record-keeping institutions (like banks and insurance companies) to remove their interests from the general privacy rights umbrella. See Regan, supra note 8, at 77–85; see also Jerry Berman & Janlori Goldman, A Federal Right of Informational Privacy: The Need for Reform (1989).
27. George Orwell, Nineteen Eighty-Four (1949).
28. For example, recall the popularity of Arthur Koestler’s Darkness at Noon and the Broadway stage adaptation by Sidney Kingsley. Arthur Koestler, Darkness at Noon, 28/367/D.A.H. (Ref. Library 1941); Sidney Kingsley, Darkness at Noon (Daphne Hardy trans., The Modern Library 1941); Sidney Kingsley, Darkness at Noon (Daphne Hardy trans., The Modern Library 1941); Sidney Kingsley, Darkness at Noon (Daphne Hardy trans., The Modern Library 1941). In popular culture, for example, consider the success of Bob Dylan’s song Subterranean Homesick Blues (critical of overzealous government); Janis Joplin’s backup band Big Brother and the Holding Company; Stills, Crosby, Nash, and Young’s song Ohio (regarding the Kent State massacre—“tin soldiers and Nixon coming”); and Francis Ford Coppola’s movie The Conversation (1974). In news media, for example, review Anne R. Field, Big Brother Inc. May Be Closer Than You Thought, Bus. Wk., Feb. 9, 1987, at 84. In scholarly literature see, for example, John Shattuck, In the Shadow of 1984: National Identification Systems, Computer-Matching, and Privacy in the United States, 35 Hastings L.J. 992 (1984). See also Regan, supra note 8, at 81 (providing references to Big Brother rhetoric that peppered floor debates over privacy policy in both chambers of Congress).
individuals, organizations, and society as a whole. The report emphasized this concern for balancing power, and for limiting the power of state and large institutions over individuals by warning that “the net effect of computerization is that it is becoming much easier for record-keeping systems to affect people than for people to affect record-keeping systems.” Further, “[a]lthough there is nothing inherently unfair in trading some measure of privacy for a benefit, both parties to the exchange should participate in setting the terms.” The lasting legacy of the report and its Code of Fair Information Practices is the need to protect privacy, at least in part, as one powerful mechanism for leveling the playing field in a game where participants have unequal starting positions.

B. Principle 2: Restricting Access to Intimate, Sensitive, or Confidential Information

This principle does not focus on who the agent of intrusion is but on the nature of information collected or disseminated—protecting privacy when information in question meets societal standards of intimacy, sensitivity, or confidentiality. Capturing the notion that people are entitled to their secrets, this principle finds robust support in scholarship developed from a variety of disciplinary perspectives, is well entrenched in practical arenas of policy and law, and is frequently raised in privacy deliberations in public or popular arenas. Several prominent philosophical and other theoretical works on privacy hold the degree of sensitivity of information to be the key factor in determining whether a privacy violation has occurred or not. These works seek to refine the category of so-called “sensitive information” and explain why the sensitivity of information is critical in defending privacy against countervailing claims.


30. RIGHTS OF CITIZENS, supra note 29.

31. Id.

32. See, e.g., RAYMOND WACKS, PERSONAL INFORMATION: PRIVACY AND THE LAW (1989) (devoted almost entirely to establishing the foundational definition of “sensitive information”);
In the United States legal landscape, sensitive information is accorded special recognition through a series of key privacy statutes that impose restrictions on explicitly identified categories of sensitive information. Examples include the Family Educational Rights and Privacy Act of 1974, which recognizes information about students as deserving protection; the Right to Financial Privacy Act of 1978, which accords special status to information about people’s financial holdings; the Video Privacy Protection Act of 1988, which protects against unconstrained dissemination of video rental records; and the Health Insurance Portability and Accountability Act of 1996 (HIPAA), which set a deadline for adoption of privacy rules governing health and medical information by the U.S. Department of Health and Human Services. Further, the common law recognizes a tort of privacy invasion in cases where there has been a “public disclosure of embarrassing private facts about the plaintiff” or an “intrusion . . . into [the plaintiff’s] private affairs.” Similar thoughts were expressed by Samuel D. Warren and Louis D. Brandeis, who were specifically concerned with protecting information about “the private life, habits, acts, and relations of an individual.”

C. Principle 3: Curtailing Intrusions into Spaces or Spheres Deemed Private or Personal

Behind this principle is the simple and ages-old idea of the sanctity of certain spaces or, more abstractly, places. For example, “a man’s home is his castle”—a person is sovereign in her own domain. Except when there are strong countervailing claims to the contrary, this principle apparently endorses a presumption in favor of people shielding themselves from the gaze of others when they are inside their own


35. 18 U.S.C. § 2710.
38. Warren & Brandeis, supra note 4, at 216.
private places. The Bill of Rights of the U.S. Constitution expresses commitment of a protected private zone in the Third and Fourth Amendments, defining explicit limits on government access to a home—quartering soldiers in the Third, and security against search and seizure in the Fourth. The Fourth Amendment, particularly, has been featured in countless cases where privacy is judged to have been violated by law enforcement agents who have breached private zones.\footnote{40. See generally Richard C. Turkington & Anita L. Allen, Privacy Law: Cases and Materials (2d ed. 2002) (providing a discussion that specifically focuses on information and information technology); W.R. LaFave, Search and Seizure: A Treatise on the Fourth Amendment (3d ed. 1996) (providing a general discussion of Fourth Amendment cases); Daniel J. Solove & Marc Rotenberg, Information Privacy Law (2003) (providing a discussion that specifically focuses on information and information technology).}

Warren and Brandeis give rousing voice to this principle: “The common law has always recognized a man’s house as his castle, impregnable, often, even to its own officers engaged in the execution of its commands. Shall the courts thus close the front entrance to constituted authority, and open wide the back door to idle or prurient curiosity?”\footnote{41. Warren & Brandeis, supra note 4, at 90.} Warren and Brandeis, thus, endorse the principled sanctity of a private domain—in this case, the home—whether against the prying of government agents or any others.

Although in many cases Principles 2 and 3 can apply simultaneously, they are independent. In the cases of a peeping Tom, for example, spying on someone in her bedroom, or a wiretap connected to a person’s telephone, we would judge privacy violated according to Principle 3, even if only mundane or impersonal information is gathered and hence Principle 2 is not violated. A similar distinction is found in numerous legal cases involving the Fourth Amendment and, of all things, garbage. Bearing most directly on the point here is the consistent finding that people cannot claim a privacy right in their garbage unless the garbage is placed within recognized private spaces (or the “curtilage”). In \textit{California v. Greenwood},\footnote{42. 486 U.S. 35 (1988).} for example, a case that has served as precedent in many that followed, the U.S. Supreme Court concluded: “[a]ccordingly, having deposited their garbage in an area particularly suited for public inspection and, in a manner of speaking, public consumption, for the express purpose of having strangers take it,
respondents could have had no reasonable expectation of privacy in the inculpatory items that they discarded.\footnote{Id. at 37; see also LAFAVE, supra note 40, at 603.}

In insisting that privacy interests in garbage are a function not of content or constitution, but of location—whether inside or outside what is considered a person’s private sphere—courts are, in effect, finding that Principle 3 is relevant to these cases, but not Principle 2; they are not finding contents of garbage to be inherently sensitive or private information.

\textbf{D. Applying the Three Principles—Some Gray Areas}

In claiming the three-principle framework has ascended to dominance in public deliberations over privacy, I maintain that it serves as a benchmark for settling disputes, but not that the outcome of disputes, or the application of the principles, is always obvious or clear. Even when it is clear which of the three principles is relevant, it may not always be obvious precisely how to draw the relevant lines to determine whether or not that principle applies, particularly with precedent setting cases involving new applications of information technology.

We have experienced this in a number of controversial government initiatives following the September 11, 2001, terrorist attacks. The USA PATRIOT Act\footnote{Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT) Act of 2001, Pub. L. No. 107-56, 115 Stat. 272.} is one example among several where government agents have clashed with citizen advocacy organizations over attempts to redraw the boundaries of access into citizens’ private lives. Even before the September 11 attacks, however, similar disagreements persisted over deployment of Carnivore, a surveillance tool for traffic flowing through the Internet.\footnote{The FBI developed the Carnivore software, which is now typically called DCS 1000.} Although a detailed account of these cases would require too great a detour from the central arguments of this Article, both are examples of disputes in which governmental interventions are asserted and contested. There is little doubt, in other words, that Principle 1 is of central relevance; what is disputed is whether the proposals in question—greater latitude for governmental surveillance both online and off-line—abide by or violate it.

Drawing lines in the case of intimate and sensitive information is also difficult and can be controversial. For example, an open question remains on whether to designate credit headers, which contain
information such as names, addresses, phone numbers, and Social Security numbers, as “personal” or not. The Individual Reference Services Group, an industry association of information brokers, maintains they are not, while the Federal Trade Commission argues they are. Case 1, raising the question whether public records ought to be available online, provokes similar questions about court records in general, and more particularly, whether some of the information contained in them and other public records should be reclassified as personal and deserving of greater protection.46 These lines are neither static nor universal as demonstrated by the case of information about students, including grades. The Family Educational and Privacy Act of 197447 marked a switch in conventional assumptions about student records. Among other things, it prohibited disclosure of information such as performance and staff recommendations without explicit permission of the students or their parents.

Similar line-drawing controversies challenge Principle 3. Interpretations of what counts as a private space may vary across times, societies, and cultures. The case of wiretaps in the United States illustrates variability across time: in 1928, in Olmstead v. United States,48 the U.S. Supreme Court ruled that wiretapping did not constitute a breach of private space.49 By 1967, however, in what is understood as an overturning of that ruling, in Katz v. United States50 the Court concluded that tapping a person’s phone does constitute an unacceptable intrusion into inviolate space.51 At least one change this shift reflects is a change in belief about what constitutes a person’s private sphere.

The Kyllo v. United States52 decision reflects similar conflicting intuitions and opinions about what constitutes an intrusion into private space. In Kyllo, the question was whether the police’s use of a thermal

49. Id. at 466.
51. Id. at 359.
52. 533 U.S. 27 (2001).
imaging device to detect patterns of heat inside the suspect’s home—for purposes of determining whether he was growing marijuana—constituted a violation of the private sphere.\footnote{Id at 27.} In a split (five to four) ruling, the Court determined that the police were at fault for not first obtaining a warrant.\footnote{Id at 28.} Against the argument proffered by the police that use of a thermal imaging device did not constitute intrusion into physical space, Justice Scalia, writing for the majority, concluded that “[i]n the home . . . all details are intimate details, because the entire area is held safe from prying government eyes.”\footnote{Id at 37.} Quoting precedent, Justice Scalia further emphasized that the law “‘draws a firm line at the entrance to the house.’”\footnote{Id (quoting Payton v. New York, 445 U.S. 573, 590 (1980)).}

Another regularly contested area, though the preponderance of opinion seems to have shifted over time, is online privacy in the workplace. Where previously this “space” was considered personal and inviolate, recent public opinion as well as court decisions suggest that ownership of servers by business organizations trumps claims by employees that the realms of the computer systems with which they work be considered a personal sphere.\footnote{For a comprehensive overview of this area of law and news media, see the Workplace Privacy webpage of the Electronic Privacy Information Center at http://www.epic.org/privacy/workplace (last visited Jan. 17, 2004).} This shift in presumption means that employers may routinely monitor e-mails and web-surfing behaviors of their employees.\footnote{But cf. Julie Cohen, Examined Lives: Informational Privacy and the Subject as Object, 52 STAN. L. REV. 1373 (2000) (providing a more pessimistic interpretation—that the increased presence of thermal imaging and similar technologies of surveillance augurs the collapse of a protected private sphere).}

\subsection*{E. The Three Principles and Public Surveillance}

The challenge posed by public surveillance is different from that posed by cases falling within the gray areas described above. In the latter, the difficulty is drawing a line; in the former, it is falling completely outside the scope of a normative model defined by the three principles. Like many of the hard cases, public surveillance typically involves a new technology, or a newly developed application of entrenched technology that expands the capacity to observe people;
gather information about them; and process, analyze, retrieve, and disseminate it. Unlike those cases, however, public surveillance does not involve government agents seeking to expand access to citizens; or collection or disclosure of sensitive, confidential, or personal information; or intrusion into spaces or spheres normally judged to be private or personal. Although public records are initially created by government agencies, the issue of placing them online does raise troubling questions of governmental overreaching and, by definition the records are public by virtue of not falling into categories of sensitive or confidential. Tracking by radio frequency identification, similarly, would not occur in places deemed private to the subjects of tracking. Online profiling is troubling, even when the information gathered is not sensitive (excludes credit card information, for example) and when it takes place on the public Web. 59 According to the framework, therefore, it seems that public surveillance is determined not to be a privacy problem. Because this conclusion is at odds with the intuition and judgment of many people, it warrants more than simple dismissal. In this disparity lie the grounds for questioning the three-principle framework as a universal standard for public deliberations over privacy.

Before presenting an alternative, contextualized approach in the next part, one conservative response to the problem of public surveillance deserves mention. Instead of simply dismissing popular aversion to public surveillance as misguided, unfounded, or irrational, this conservative view distinguishes between privacy—the value, which is embodied in the three principles, and privacy—the more encompassing category of preference, or taste, revealed in results of numerous public opinion surveys. 60 Designating public surveillance as a member of the second category still affords it various means of social protection, in addition to “self-help.” 61 As commonly understood, democratic market-based societies offer at least two robust mechanisms for expressing popular preference: first, citizens can press for laws to protect majority

59. The term “public Web” is used to mark a distinction between those realms of the Web that are publicly accessible and those that are accessible only to authorized users and frequently protected by some form of security.


preferences, and second, consumers, through their actions, can affect the
terms and nature of commercial offerings in a free, competitive
marketplace. These alternatives deserve a great deal more attention
than I am able to offer here.

Although this view preserves the three-principle framework, at least
one problem with it is that it places resistance to public surveillance on a
weak footing against countervailing claims, particularly those backed by
recognized rights and values. In a free society, a person has a right to
choose chocolate over vanilla ice cream, or to press for extensive
protections of privacy preferences, except where such preferences
happen to conflict with another person’s claim to something of greater
moral or political standing. Those who conduct public surveillance, or
support its pursuit, have lobbied exactly on those grounds, citing such
well-entrenched freedoms as speech, action, and pursuit of wealth. The
weak footing that this allows for the aversion to public surveillance can
be demonstrated in relation to a commonly used legal standard, namely,
reasonable expectation of privacy.

Justice John Harlan, concurring with the majority opinion in Katz, is
credited with formulating two conditions that later courts have used to
test whether a person has “a reasonable expectation of privacy” in any
given activity or practice, namely: (1) that the person exhibited an actual
expectation of privacy, and (2) that the expectation is one that society is
prepared to recognize as reasonable. Although the reasonable
expectation benchmark raises deep and complex questions that cannot be
addressed here, there is at least one point of direct interest, notably that
the benchmark is a potential source of crushing rebuttal to preference-
based complaints against public surveillance. It is simply this: when
people move about and do things in public arenas, they have implicitly
yielded any expectation of privacy. Much as they might prefer that
others neither see, nor take note, expecting others not to see, notice, or

62. Privacy skeptics have argued that because people seem to do neither, they obviously do not
care much about privacy. See Calvin C. Gotlieb, Privacy: A Concept Whose Time Has Come and
Gone, in COMPUTERS, SURVEILLANCE, AND PRIVACY 156 (David Lyon & Elia Zureik eds., 1996);
Solveig Singleton, Privacy as Censorship: A Skeptical View of Proposals To Regulate Privacy in
the Private Sector, in CATO POL’Y ANALYSIS NO. 295, (Cato Inst. 1998), available at
63. Many articles deal with privacy in relation to competing claims. But see, e.g., Cohen, supra
note 58, at 1373; Richard Posner, The Right to Privacy, 12 GA. L. REV. 393 (1978); Eugene Volokh,
Personalization and Privacy, COMM. ACM, Aug. 2000, at 84.
64. See Katz v. United States, 389 U.S. 347, 360–61 (1967) (Harlan, J., concurring); see also
REGAN, supra note 8, at 122; SOLOVE & ROTENBERG, supra note 40, at 21.
make use of information so gained would be unreasonably restrictive of others’ freedoms. One cannot reasonably insist that people avert their eyes, not look out their windows, or not notice what others have placed in their supermarket trolleys. And if we cannot stop them from looking, we cannot stop them remembering and telling others. In 2001, Tampa police, defending their use of video cameras to scan faces one-by-one as they entered the Super Bowl stadium, stated, “the courts have ruled that there is no expectation of privacy in a public setting.”

In sum, maintaining that the three principles define the value of privacy provides significant force to the reasonableness of privacy claims covered by them, but offers little cover for anything outside the principles. Cast as preference, these claims are not ruled out as grounds for favoring one outcome over another, though not accorded special consideration in competition with others. Accordingly, there is no prima facie concern over placing public records, already available for anyone to see, online, or for permitting aggregation of non-sensitive information, so long as a compelling reason such as efficiency, safety, or profit can be offered. Since RFID and other surveillance are conducted in public venues only, the expectation of privacy in any of these contexts cannot be reasonable. Those who hold that public surveillance can constitute a violation and not merely a practice that some people dislike will remain unconvinced.

III. CONTEXTUAL INTEGRITY

Highlighting two features of the three-principle framework helps to convey what lies behind the idea of contextual integrity. One is that it is posed as a universal account of what does and does not warrant restrictive, privacy-motivated measures. That is, as a conceptual framework, it is not conditioned on dimensions of time, location, and so forth. Another is that it expresses a right to privacy in terms of dichotomies—sensitive and non-sensitive, private and public, government and private—that line up, interestingly, with aspects of the general public-private dichotomy that has been useful in other areas of political and legal inquiry. That which falls within any one of the appropriate halves warrants privacy consideration; for all the rest,

66. It might still admit of variability in that the categories of sensitive and non-sensitive, for example, could vary across, say, cultures, historical periods, and places.
Privacy as Contextual Integrity

anything goes. In both these features, the account of privacy in terms of contextual integrity diverges from the three-principle model.

A central tenet of contextual integrity is that there are no arenas of life not governed by norms of information flow, no information or spheres of life for which “anything goes.” Almost everything—things that we do, events that occur, transactions that take place—happens in a context not only of place but of politics, convention, and cultural expectation. These contexts can be as sweepingly defined as, say, spheres of life such as education, politics, and the marketplace or as finely drawn as the conventional routines of visiting the dentist, attending a family wedding, or interviewing for a job. For some purposes, broad sweeps are sufficient. As mentioned before, public and private define a dichotomy of spheres that have proven useful in legal and political inquiry. Robust intuitions about privacy norms, however, seem to be rooted in the details of rather more limited contexts, spheres, or stereotypic situations.

Observing the texture of people’s lives, we find them not only crossing dichotomies, but moving about, into, and out of a plurality of distinct realms. They are at home with families, they go to work, they seek medical care, visit friends, consult with psychiatrists, talk with lawyers, go to the bank, attend religious services, vote, shop, and more. Each of these spheres, realms, or contexts involves, indeed may even be defined by, a distinct set of norms, which governs its various aspects such as roles, expectations, actions, and practices. For certain contexts, such as the highly ritualized settings of many church services, these norms are explicit and quite specific. For others, the norms may be implicit, variable, and incomplete (or partial). There is no need here to construct a theory of these contexts. It is enough for our purposes that the social phenomenon of distinct types of contexts, domains, spheres, institutions, or fields is firmly rooted in common experience and has been theorized in the profound work of reputable philosophers, social scientists, and social theorists.67 Any of these sources could provide

67. See generally PIERRE BOURDIEU & LOIC J.D. WACQUANT, AN INVITATION TO REFLEXIVE SOCIOLOGY 95–115 (1992) (providing general discussion of Pierre Bourdieu’s fields); id. at 97 (“In highly differentiated societies, the social cosmos is made up of a number of such relatively autonomous social microcosms . . . . For instance, the artistic field, or the religious field, or the economic field all follow specific logics . . . .”); MICHAEL PHILLIPS, BETWEEN UNIVERSALISM AND SKEPTICISM: ETHICS AS SOCIAL ARTIFACT (1994); MICHAEL WALZER, SPHERES OF JUSTICE: A DEFENSE OF PLURALISM AND EQUALITY (1983); Roger Friedland & Robert R. Alford, Bringing Society Back In: Symbolic Practices, and Institutional Contradictions, in THE NEW INSTITUTIONALISM IN ORGANIZATIONAL ANALYSIS 232, 247–59 (Walter W. Powell & Paul J. DiMaggio eds., 1991) (also discussing institutions); id. at 251 (“[Institutions] generate not only that
foundational concepts for articulating the concept of contextual integrity in relation to personal information.

Contexts, or spheres, offer a platform for a normative account of privacy in terms of contextual integrity. As mentioned before, contexts are partly constituted by norms, which determine and govern key aspects such as roles, expectations, behaviors, and limits. There are numerous possible sources of contextual norms, including history, culture, law, convention, etc. Among the norms present in most contexts are ones that govern information, and, most relevant to our discussion, information about the people involved in the contexts. I posit two types of informational norms: norms of appropriateness, and norms of flow or distribution. Contextual integrity is maintained when both types of norms are upheld, and it is violated when either of the norms is violated. The central thesis of this Article is that the benchmark of privacy is contextual integrity; that in any given situation, a complaint that privacy has been violated is sound in the event that one or the other types of the informational norms has been transgressed.68

A. Appropriateness

As the label suggests, norms of appropriateness dictate what information about persons is appropriate, or fitting, to reveal in a particular context. Generally, these norms circumscribe the type or nature of information about various individuals that, within a given context, is allowable, expected, or even demanded to be revealed. In medical contexts, it is appropriate to share details of our physical condition or, more specifically, the patient shares information about his or her physical condition with the physician but not vice versa; among friends we may pour over romantic entanglements (our own and those of others); to the bank or our creditors, we reveal financial information; with our professors, we discuss our own grades; at work, it is appropriate to discuss work-related goals and the details and quality of performance.

As important is what is not appropriate: we are not (at least in the United States) expected to share our religious affiliation with employers,
Privacy as Contextual Integrity

financial standing with friends and acquaintances, performance at work with physicians, etc. As with other defining aspects of contexts and spheres, there can be great variability from one context to the next in terms of how restrictive, explicit, and complete the norms of appropriateness are. In the context of friendship, for example, norms are quite open-ended, less so in the context of, say, a classroom, and even less so in a courtroom, where norms of appropriateness regulate almost every piece of information presented to it. The point to note is that there is no place not governed by at least some informational norms. The notion that when individuals venture out in public—a street, a square, a park, a market, a football game—no norms are in operation, that “anything goes,” is pure fiction. For example, even in the most public of places, it is not out of order for people to respond in word or thought, “none of your business,” to a stranger asking their names.

While norms of appropriateness are robust in everyday experience, the idea that such norms operate has not been explicitly addressed in most of the dominant research and scholarship that feed into public deliberations of privacy policy in the United States.69 Within the philosophical literature of the past few decades, however, we find recognition of similar notions. James Rachels, for example, has posited something like a norm of appropriateness in arguing that adequate privacy protection accords people the important power to share information discriminately, which in turn enables them to determine not only how close they are to others, but the nature of their relationships: businessman to employee, minister to congregant, doctor to patient, husband to wife, parent to child, and so on. In each case, the sort of relationship that people have to one another involves a conception of how it is appropriate for them to behave with each other, and what is more, a conception of the kind and degree of knowledge concerning one another which it is appropriate for them to have.70

Ferdinand Schoeman, a philosopher who has offered one of the deepest and most subtle accounts of privacy and its value to humans, writes,

69. The formal regulation of confidentiality within professional fields is an exception, but this Article argues that similar norms hold in all contexts, even if not stipulated in explicit laws or regulations.
[p]eople have, and it is important that they maintain, different relationships with different people.” Further,

[a] person can be active in the gay pride movement in San Francisco, but be private about her sexual preferences vis-à-vis her family and coworkers in Sacramento. A professor may be highly visible to other gays at the gay bar but discreet about sexual orientation at the university. Surely the streets and newspapers of San Francisco are public places as are the gay bars in the quiet university town. Does appearing in some public settings as a gay activist mean that the person concerned has waived her rights to civil inattention, to feeling violated if confronted in another setting? These cases illustrate Schoeman’s sense that appropriating information from one situation and inserting it in another can constitute a violation. Violations of this type are captured with the concept of appropriateness.

B. Distribution

In addition to appropriateness, another set of norms govern what I will call flow or distribution of information—movement, or transfer of information from one party to another or others. The idea that contextual norms regulate flow or distribution of information was profoundly influenced by Michael Walzer’s pluralist theory of justice. Although Walzer’s theory does not specifically address the problems of privacy and regulation of information, it provides insights that are useful to the construction of privacy as contextual integrity.

In his book, Spheres of Justice: A Defense of Pluralism, Walzer develops a theory of distributive justice in terms of not only a single good and universal equality, but in terms of something he calls complex equality, adjudicated across distinct distributive spheres, each with its own, unique set of norms of justice. Walzer conceives of societies as made up of numerous distributive spheres, each defined by a social good internal to them. Social goods include such things as wealth, political

71. Ferdinand Schoeman, Privacy and Intimate Information, in PHILOSOPHICAL DIMENSIONS OF PRIVACY: AN ANTHOLOGY, supra note 70, at 403, 408.


73. See WALZER, supra note 67. Jeroen van den Hoven pointed out the relevance of this work to me.

74. See id.

75. See generally id.
Privacy as Contextual Integrity

office, honor, commodities, education, security and welfare, and employment. These social goods are distributed according to criteria or principles that vary according to the spheres within which they operate. In the educational sphere, for example, access to instruction up to a certain level (a good) might be guaranteed to all residents of a community with appropriate mental capacities and instruction beyond the basic level, say, a university undergraduate education, allocated only to those who have performed to a particular standard. Commodities (goods) in a marketplace are distributed according to preferences and ability to pay; in the sphere of employment, jobs (goods) are allocated to those with appropriate talents and qualifications, and so on. According to Walzer, complex equality, the mark of justice, is achieved when social goods are distributed according to different standards of distribution in different spheres and the spheres are relatively autonomous. Thus, in Walzer’s just society, we would see “different outcomes for different people in different spheres.”

Complex equality adds the idea of distributive principles or distributive criteria to the notion of contextual integrity. What matters is not only whether information is appropriate or inappropriate for a given context, but whether its distribution, or flow, respects contextual norms of information flow.

Let us return to the context of friendship, this time to consider some examples of norms of flow. As described earlier, relatively few general norms of appropriateness apply, though practices may vary depending on whether the friends are close, have known each other for a long time, and so on. Information that is appropriate to friendship can include mundane information about day-to-day activities, likes and dislikes, opinions, relationships, character, emotions, capacity for loyalty, and much more. The same open-endedness, however, does not hold for norms of flow, which are quite substantial. In friendship, generally, information is either shared at the discretion of the subject in a bidirectional flow—friends choose to tell each other about themselves—or is inferred by one friend of another on the basis of what the other has done, said, experienced, etc. But that is not all. Confidentiality is

76. See generally id.
77. See generally id.
78. See generally id.
79. See generally id.
80. Id. at 320.
generally the default—that is, friends expect what they say to each other to be held in confidence and not arbitrarily spread to others. While some departure from the norms is generally allowable, as when friends coax information from each other, straying too far is usually viewed as a serious breach. Where a friend ferrets out information from third party sources, or divulges information shared in friendship to others for reasons having nothing to do with the friendship, not only might the friend justifiably feel betrayed, but the actions may call into question the very nature of the relationship.81

Free choice, discretion, and confidentiality, prominent among norms of flow in friendship, are not the only principles of information distribution. Others include need, entitlement, and obligation—a list that is probably open-ended. In a healthcare context, for example, when a patient shares with her physician details of her current and past physical condition, the reigning norm is not discretion of the subject (that is, free choice of the patient) but is closer to being mandated by the physician who might reasonably condition treatment on a patient’s readiness to share information that the physician deems necessary for competent diagnosis and treatment. Another difference from friendship is that in the healthcare context, the flow is not normally bidirectional. Confidentiality of patient health information is the subject of complex norms—in the United States, for example, a recent law stipulates when, and in what ways, a physician is bound by a patient’s consent: for example, where it is directly pertinent to diagnosis and treatment, where it poses a public health risk, and where it is of commercial interest to drug companies. 82

Other cases of information practices following rational norms of flow include, for example, transactions between customers and mail-order merchants. In such transactions, customers are required to provide sufficient and appropriate information to satisfy companies that they can pay, and provide an address indicating where packages should be sent. Police are bound by law to abide by various regulations governing modes of acquiring information and how to deal with its flow thereafter. However, suspects arrested by police on criminal charges may volunteer

81. We may wonder how it would affect a friendship if one party discovers his friend has engaged the help of the much advertised snoop programs that promise the ability to track e-mail correspondence.

Privacy as Contextual Integrity

certain categories of information beyond those that they are compelled to provide. A sexual partner may be entitled to information about the other’s HIV status, although the same demand by a friend is probably not warranted. A job applicant may volunteer information she considers evidence of her ability to do the job. Candidates for political office volunteer proof of professional competence, political loyalty, personal integrity, political connections, and past political activities. But it is accepted that employers and voters, respectively, might choose to conduct independent investigations as to fitness and competence. These cases are intended merely to illustrate the many possible configurations of informational norms we are likely to encounter, and they just begin to scratch the surface.

C. Change, Contextual Integrity, and Justice

As proposed above, a normative account of privacy in terms of contextual integrity asserts that a privacy violation has occurred when either contextual norms of appropriateness or norms of flow have been breached. One point of contrast with other theoretical accounts of privacy rights is that personal information revealed in a particular context is always tagged with that context and never “up for grabs” as other accounts would have us believe of public information or information gathered in public places. A second point of contrast is that the scope of informational norms is always internal to a given context, and, in this sense, these norms are relative, or non-universal. Before visiting the problem of public surveillance in light of contextual integrity, two potentially worrisome implications should first be addressed, both consequences of this built-in contextual dependence.

One is that by putting forward existing informational norms as benchmarks for privacy protection, we appear to endorse entrenched flows that might be deleterious even in the face of technological means to make things better. Put another way, contextual integrity is conservative in possibly detrimental ways. As a brief example, consider the substantial benefits that networked information systems with good search capabilities provide consumers wishing to find out more about products, services, or service providers, say, to check whether a particular surgeon has been found guilty of malpractice. Because the capabilities are new, ferreting out such information constitutes a radical departure from past practice, which, in the case of the surgeon, might have meant a patient having to ask the surgeon directly or engage someone else in a costly search. It would be problematic if the theory of
contextual integrity would judge new forms of information gathering to be a privacy violation in such instances.

A second worry is that contextual integrity, being so tied to practice and convention, loses prescriptive value or moral authority. In this era of rapid transformations due to computing and information technologies, changes are thrust upon people and societies frequently without the possibility of careful deliberation over potential harms and benefits, over whether we want or need them.\textsuperscript{83} Practices shift almost imperceptibly but, over time, quite dramatically, and in turn bring about shifts in conventional expectations. These changes have influenced outcomes in a number of important cases, such as determining that the Fourth Amendment was not breached when police discovered marijuana plants in a suspect’s yard by flying over in a surveillance plane.\textsuperscript{84}

The U.S. Supreme Court held that people do not have a reasonable expectation of privacy from air surveillance because flights have become a common part of our lives.\textsuperscript{85} In \textit{Kyllo},\textsuperscript{86} even though the Court concluded the Fourth Amendment had been breached, one of the reasons for its conclusion was that thermal imaging trained on a private residence (unlike plane flights) was not yet common practice and so would count as a search.\textsuperscript{87} As long as contextual integrity is tied, in these ways, to practice and convention, it would be unconvincing as a source of moral prescription, that is, constituting adequate justification for what one morally should or should not do.

Although the two worries come from apparently opposite directions, in fact, they provoke a similar set of elaborations. First, they highlight the importance of distinguishing \textit{actual} practice from \textit{prescribed} practice. Second, even within the category of prescribed practice, the grounds for prescription can vary among several possibilities. Third, even entrenched norms can change over time and may vary across not only historical moments, but cultures, geographic locations, societies, nations, etc. Although these considerations mean that just because something is the case, does not mean it morally or politically ought to be the case, they also mean that something more is needed to enable us to

\textsuperscript{83} I am aware of an oversimplification in the way I express this issue, for change is not strictly a consequence of devices and systems by themselves but, of course, may involve other social, economic, or legal determinants.


\textsuperscript{85} \textit{Id.} at 458.

\textsuperscript{86} See supra notes 52–56 and accompanying text.

\textsuperscript{87} See \textit{Kyllo} v. United States, 533 U.S. 27, 34, 40 (2001).
Privacy as Contextual Integrity

distinguish changes that are morally and politically acceptable, or even desirable, from those that are not (and ought to be resisted). As explained below, this can be done, but only indirectly.

I propose that the requirement of contextual integrity sets up a presumption in favor of the status quo; common practices are understood to reflect norms of appropriateness and flow, and breaches of these norms are held to be violations of privacy. Walzer's account of justice asserts a similar presumption in the case of spheres, namely, that distributing social goods of one sphere according to criteria of another constitutes injustice.88 Evidence of a commitment to this presumption is that our society recognizes as wrong wealthy people buying favorable verdicts in courts of law, bosses demanding sexual favors as a condition of promotion, awarding political office on the basis of kinship, and determining wage scales by gender or race. These examples are unjust not only because goods from one sphere have intruded into another, but also because distributional norms of one sphere are being applied to another. Further, Walzer considers it a form of tyranny when goods of one sphere intrude into, or become dominant in, not only one sphere but many; local norms that embody the settled rationale of the tyrannized sphere are overturned as those who possess vast amounts of dominant goods are able to exert tyrannical power over those who do not.89

A presumption in favor of the status quo for informational norms means we initially resist breaches, suspicious that they occasion injustice or even tyranny. We take the stance that the entrenched normative framework represents a settled rationale for a certain context that we ought to protect unless powerful reasons support change. The settled rationale of any given context may have long historical roots and serve important cultural, social, and personal ends. The hugely complex system of regulations in the medical context can be traced at least as far back as the fourth century B.C.E., when Hippocrates exhorted fellow physicians to maintain confidentiality because of the shame involved in passing on any further what they learn about their patients in the course of treatment: "And about whatever I may see or hear in treatment, or even without treatment, in the life of human beings—things that should

89. Id.
not ever be blurted out outside—I will remain silent, holding such things
to be unutterable [sacred, not to be divulged].90

The context of elections for political office is another case of a settled
normative framework that functions in generally positive ways.91 On
election day, citizens converge on polling stations to cast votes. From the
moment they cross the threshold, information flows are highly regulated,
from what elections officers can ask them to what they can ask officers,
what voters are required to document in writing, who sees it, what
happens to the vote cast and who sees that, what exit pollsters can ask
citizens as they leave—for whom they voted but not voters’ names—and
what the exit pollsters are free to disseminate publicly. These two
familiar cases illustrate how systems of norms of appropriateness and
flow may evolve to serve determinable ends and institutions.

A presumption in favor of status quo does not, however, rule out the
possibility of a successful challenge where adequate reasons exist.
Resolving these contested cases calls for reliable means of evaluating the
relative moral standing of entrenched norms and the novel practices that
breach or threaten them. Specifically, I propose that entrenched norms be
compared with novel practices that breach or threaten them, and judged
worth preserving, or not, in terms of how well they promote not only
values and goods internal to a given context, but also fundamental social,
political, and moral values. Conducting the second of these two modes
of evaluation, namely, a comparison in terms of social, political, and
moral values, involves identifying fundamental values that may be
served by (or obscured by) the relevant informational norms imposing
restrictions on the flow and distribution of personal information in the
given case. According to the insights of several privacy scholars, the list
of values likely to be affected includes: (1) prevention of information-
based harm, (2) informational inequality, (3) autonomy, (4) freedom,
(5) preservation of important human relationships, and (6) democracy
and other social values.92 Values that are regularly cited in support of

90. "In a Pure and Holy Way:” Personal and Professional Conduct in the Hippocratic Oath, 51

91. I am speaking of elections in a democratic state, with details drawn more specifically from the
case of the United States.

92. This list is informed by the work of Julie Cohen, Stanley Benn, Ruth Gavison, Jeroen van den
Hoven, James Nehf, Paul Schwartz, Jeffrey Reiman, Jeffrey Rosen, and others. Citations to specific
works are given in footnotes to follow.
free or unconstrained flows include: (1) freedom of speech, (2) pursuit of wealth, (3) efficiency, and (4) security.

1. Prevention of Informational Harms

Information in the wrong hands or generally unrestricted access to information can be harmful. The harm in question can be severe, such as occurred in the case of the murder of actress Rebecca Schaeffer in 1989, when it was discovered that the murderer located her home address through Department of Motor Vehicles records.93 Less palpable, but also serious, are harms like identity theft, which occurs with increasing frequency, apparently as a result of the ready availability of key identifying information like Social Security numbers, addresses, and phone numbers. Furthermore, various goods such as employment, life, and medical insurance, could be placed at risk if the flow of medical information were not restricted, or if information regarding people’s religious and political affiliations, sexual orientation, or criminal records were readily available.

2. Informational Inequality

There are a number of facets to this value. In the crucial 1973 U.S. Department of Health, Education, and Welfare’s report on computerized records, the opening sentences presented fairness, or we might say justice, as a foundational value for regulating the collection, storage, and use of personal information in computerized databases.94 The Department’s politically grounded argument will be familiar in the American contexts where entities, such as government and financial institutions, wield significant power over the fates of individual citizens and clients. Allowing these institutions free reign in collecting and using information further tips the balance of power in their favor. Responsive to the strong sentiment in favor of leveling the playing field, the widely influential Code of Fair Information Practices defined restrictions on gathering, storing, and using information about people in the name of fairness.95

---

94. RIGHTS OF CITIZENS, supra note 29.
95. Id. at xxiii–xxxv.
Inequalities may also arise in the context of routine commercial transactions mediated by technologies of information. As described by Jeroen van den Hoven, individuals acquiring goods or services are also giving (some would say, selling) something, namely, information about themselves, such as their credit card numbers, names, or addresses.\textsuperscript{96} Usually the parties in the transaction are far from equal. For the most part, individuals have little knowledge and understanding of the potential value of this economic exchange; do not know what will be done with the information; do not grasp the full implications of consenting to release of information; and almost certainly have no power to retract or redraw the arrangement should it prove annoying, burdensome, or simply different from what they had initially sought. van den Hoven calls for “openness, transparency, participation, and notification on the part of business firms and direct marketers to secure fair contracts,” in order to promote fairness in exchange.\textsuperscript{97}

3. Autonomy and Freedom

For purposes of this abbreviated discussion, we consider ways in which autonomy and freedom, taken together, have indicated the need for wise restrictions on access to personal information.\textsuperscript{98} Typically associated with the liberal political vision, autonomy is the mark of thoughtful citizens whose lives and choices are guided by principles they have adopted as a result of critical reflection.\textsuperscript{99} Thoughtful works on privacy by Ruth Gavison, Jeffrey Reiman, Julie Cohen, and others have demonstrated a rich array of associations between autonomy and privacy.\textsuperscript{100} These works assert that freedom from scrutiny and zones of “relative insularity”\textsuperscript{101} are necessary conditions for formulating goals, values, conceptions of self, and principles of action because they provide venues in which people are free to experiment, act, and decide without

\textsuperscript{96} van den Hoven, supra note 67.
\textsuperscript{97} Id. at 435.
\textsuperscript{98} Consider the title of Alan F. Westin’s early and influential book, Privacy and Freedom. ALAN F. WESTIN, PRIVACY AND FREEDOM (1967).
\textsuperscript{101} See, e.g., Cohen, supra note 58, at 1424.
Privacy as Contextual Integrity

giving account to others or being fearful of retribution. Uninhibited by what others might say, how they will react, and how they will judge, unhindered by the constraints and expectations of tradition and convention, people are freer to formulate for themselves the reasons behind significant life choices, preferences, and commitments. In defending robust broad protections for informational privacy, Cohen reminds us that autonomy touches many dimensions of peoples’ lives, including tastes, behaviors, beliefs, preferences, moral commitments, associations, decisions, and choices that define who we are.

Besides the causal or enabling connection between privacy and autonomy, a further, constitutive connection that is hardly ever recognized as such plays an essential role in the most widely held definition of a right to privacy—the right to control information about oneself. The plausibility of such a right to control information about oneself, even one that is limited and constrained by other competing or countervailing rights and obligations, rests on the premise that information about ourselves is something over which individuals may exercise autonomy. In this way, it is comparable to the prima facie rights of self-determination that we have over our bodies and access to them.

4. Preservation of Important Human Relationships

Information is a key factor in the relationships we have and form with others. Charles Fried has said that controlling who has access to personal information about ourselves is a necessary condition for friendship, intimacy, and trust. James Rachels, as mentioned earlier, has made a related point that distinctive relationships, for example individual to spouse, boss, friend, colleague, priest, teacher, therapist, hairdresser, and so on, are partially defined by distinctive patterns of information sharing. Insofar as these relationships are valued, so would we value adequate and appropriate restrictions on information flows that bolster them.

---

102. See, e.g., Gavison, supra note 10.
103. Cohen, supra note 58, at 1425.
104. See REGAN, supra note 8; WESTIN, supra note 98; Cohen, supra note 58. However, I believe this conception is deeply flawed for reasons offered by Ruth Gavison and Jeffrey Reiman. See Gavison, supra note 10; Reiman, supra note 100.
105. See Fried, supra note 32.
106. See Rachels, supra note 70.
5. Democracy and Other Social Values

Several proponents of strong privacy protections point out the importance of privacy not only to individuals but to society. Priscilla Regan, in *Legislating Privacy*, provides one of the best-informed versions of this claim:

Privacy has value beyond its usefulness in helping the individual maintain his or her dignity or develop personal relationships. Most privacy scholars emphasize that the individual is better off if privacy exists; I argue that society is better off as well when privacy exists. I maintain that privacy serves not just individual interests but also common, public, and collective purposes.107

Regan and others describe ways in which privacy is essential to nourishing and promoting the values of a liberal, democratic, political, and social order by arguing that the vitality of democracy depends not only on an autonomous and thoughtful citizenry—bolstered through privacy—but on the concrete protection against public scrutiny of certain spheres of decision-making, including but not limited to the voting booth.108 Privacy is a necessary condition for construction of what Erving Goffman calls “social personae,” which serves not only to alleviate complex role demands on individuals, but to facilitate a smoother transactional space for the many routine interactions that contribute to social welfare.109 Similar arguments have been offered by Janlori Goldman defending robust protections of medical information on grounds that individuals would then be more likely both to seek medical care and agree to participate in medical research. In turn, this would improve overall public health as well as social welfare through scientific research. Arguments favoring restrictions of online transactional information cite potential gains, namely, the increased likelihood of participation in electronic commerce.110 Finally, Oscar Gandy has

107. REGAN, supra note 8, at 221.
108. See Anita L. Allen, Coercing Privacy, 40 WM. & MARY L. REV. 723 (1999); Cohen, supra note 58; Janlori Goldman, Protecting Privacy to Improve Health Care, HEALTH CARE, Nov./Dec. 1998, at 47.
109. See ERVING GOFFMAN, THE PRESENTATION OF SELF IN EVERYDAY LIFE (1959) (discussing the importance of maintaining a “backstage” where people are allowed to relax out of character (ch. 3) and describing the preferences of both audiences and performers to maintain a façade in various ritualized social settings, even when both know that the performances in question do not reveal the whole truth (ch. 6)); see also Cohen, supra note 58, at 1427.
Privacy as Contextual Integrity

vividly conveyed how profiling and the widespread collection, aggregation, and mining of data increase social injustice and generate even further discrimination against traditionally disadvantaged ethnic groups. 111

6. Countervailing Values

There are obviously many reasons for favoring the collection, sharing, and widespread distribution of personal information, including maintaining free speech 112 and a free press, economic efficiency 113 and profitability, open government, and security. 114 When these values clash with those that support restrictive treatment, we need to pursue trade-offs and balance.

D. Applying Contextual Integrity to the Three Cases

One of the key ways contextual integrity differs from other theoretical approaches to privacy is that it recognizes a richer, more comprehensive set of relevant parameters. In addressing whether placing public records online is problematic, whether moving records from filing cabinets or stand-alone databases onto the net marks a significant change, it forces us to look beyond whether the information in question is public. To establish whether contextual integrity is breached requires an examination of governing norms of appropriateness and flow to see whether and in what ways the proposed new practices measure up.

When the first case, the availability of public records online, is viewed through the lens of contextual integrity, certain aspects of the

111. See Oscar H. Gandy, Jr., THE PANOPTIC SORT: A POLITICAL ECONOMY OF PERSONAL INFORMATION (1993); Oscar H. Gandy, Jr., Coming to Terms with the Panoptic Sort, in COMPUTERS, SURVEILLANCE, AND PRIVACY 132 (David Lyon & Elia Zureik eds. 1996); Oscar H. Gandy, Jr., Exploring Identity and Identification, 14 NOTRE DAME J.L. ETHICS & PUB. POL’Y 1085 (2000).

112. See, e.g., Cohen, supra note 58; Paul M. Schwartz, Privacy and Democracy in Cyberspace, 52 VAND. L. REV. 1607 (1999); Volokh, supra note 63, at 84; see also SOLOVE & ROTENBERG, supra note 40, ch. 2, sec. C (providing extensive case law).

113. See Singleton, supra note 62 (describing economic efficiency as potentially in conflict with privacy).

114. See Orrin Kerr, Internet Surveillance Law After the USA PATRIOT Act: The Big Brother That Isn’t, 97 NW. U. L. REV. 607 (2003). In general, literature and cases surrounding the Fourth Amendment involve a quest to balance privacy against security.
change in placement from locally kept records (whether hardcopy or electronic) to Web-accessible records, are highlighted in novel ways. The change in placement, which vastly alters the range of accessibility from local to global, is significant because it constitutes a breach of entrenched norms of flow. As such, it demands scrutiny in terms of values. Although a full-blown analysis is not possible in the context of this Article, it is instructive to consider, briefly, how this affects a case that, arguably, draws little sympathy—the convicted sex offender. Recent changes in the laws of various states require that neighbors be informed if someone with a record of a serious sex offense moves into the neighborhood.115 Despite objections, a good case may be made in favor of altering the distributional norms, from storing a record in a publicly available cabinet to actively informing neighbors. A proposal to place these records online, however, is different. While residents of, say, Hamilton, New Jersey, might reasonably argue that being informed about a released sex offender in their neighborhood is a justified measure of protection against the dangers of recidivism, believed to be high in the case of sex crimes, a similar argument seems specious for a citizen of, say, Fairbanks, Alaska. Furthermore, placing the myriad categories of public records online would greatly facilitate the aggregation and analysis of these records by third parties. This radical alteration of availability and flow does little to address the original basis for creation of public records, namely, public accountability of governmental agencies.116

The second case, consumer profiling and data mining, can be analyzed in a similar way. As before, the crucial issue is not whether the information is private or public, gathered from private or public settings, but whether the action breaches contextual integrity. The use of credit cards and the emergence of information brokers, along with a host of technical developments, however, have altered patterns of availability and flow in well-known ways. But are these changes significant from the perspective of contextual integrity? The answer is variable. In the past, it was integral to the transaction between a merchant and a customer that the merchant would get to know what a customer purchased. Good, that is to say, competent merchants, paying attention to what customers wanted, would provide stock accordingly. Although the online bookseller Amazon.com maintains and analyzes customer records

116. REPORT OF THE SPECIAL DIRECTIVE SUBCOMMITTEE, supra note 46; Gellman, supra note 2.
Privacy as Contextual Integrity

electronically, using this information as a basis for marketing to those same customers seems not to be a significant departure from entrenched norms of appropriateness and flow. By contrast, the grocer who bombards shoppers with questions about other lifestyle choices—e.g., where they vacationed, what movies they recently viewed, what books they read, where their children attend school or college, and so on—does breach norms of appropriateness. The grocer who provides information about grocery purchases to vendors of magazine subscriptions or information brokers like Seisint and Axiom is responsible not only for breaches of norms of appropriateness but also norms of flow.¹¹⁷

Contextual integrity generates similar questions about RFID tags because they too significantly alter the nature and distribution patterns of information. Prior to the advent of RFID tags, customers could assume that sales assistants, store managers, or company leaders recorded point-of-sale information. RFID tags extend the duration of the relationships, making available to the jeans retailer, the manufacturer, and others a range of information about customers that was not previously available. These potential uses of RFID tags can affect not only who gains access to customer information, but at whose discretion. In a departure from past assumption, the customer would no longer control the distribution of information beyond point of sale. Unless RFID tags are designed specifically to allow for easy detection and disabling, discretion is removed from the customer and placed into the hands of information gatherers. This departure from entrenched norms triggers an assessment in terms of values.

E. Contextual Integrity and Other Privacy-Centric Approaches

For the three cases, I have been able to provide only sketches of arguments to support particular prescriptions to restrict (or not restrict) information gathering, aggregation, and dissemination on the basis of contextual integrity. In general, the norms of appropriateness and flow demand consideration of a number of parameters, including the nature of the information in question and its relationship to the context, the roles involved in the context, the relationships among the roles, the rules of flow, and how any changes made within a context might affect the underlying values. For the most part, building a conclusive argument in

¹¹⁷. To complete the argument would require showing that these breaches are justifiable neither in terms of values internal to the context nor in terms of more fundamental social, political, and moral values.
terms of contextual integrity involves painstaking analysis of details (or building upon analyses of identical or very similar cases), including even a reference to factual findings, which might ground claims about the empirical effects of a change on key parameters.

In developing the rationale for a new way of thinking about some of the puzzles of public surveillance or “privacy in public,”118 deficiencies (or blind spots) in the three-principle framework served as a springboard for an alternative normative theory built around the concept of contextual integrity. Although this strategy highlights the specific strength of contextual integrity to resolve puzzles of public surveillance, it gives short shrift to a body of theoretical works on privacy—many proposed in the past few years—whose broadly encompassing privacy principles also extend to various forms of public surveillance, among other things.119 Given space constraints, I am not able here to give them the degree of individual consideration they deserve except briefly to mention the one most significant point of contrast. Where these other accounts offer interpretations of privacy in terms of universal prescriptions, contextual integrity couches its prescriptions always within the bounds of a given context.

The widely held conception of a right of privacy as a right to control information about oneself, for example, is sufficiently capacious to entail protections even in categories of so-called public information, public spaces, and against non-governmental agents. The same potential holds for rights posited in terms of freedom from visual surveillance or restrictions on access to the subject. From the perspective of contextual integrity, where prescriptions are always couched in context-specific terms, these conceptions would be considered too blunt, possibly dogmatic. Even allowing for tradeoffs with other competing claims and rights, for balancing privacy against other values such as security, property, or speech (which any reasonable version would), the claim to control and limit access remains too open-ended and still leaves out too much of the picture.

According to the theory of contextual integrity, it is crucial to know the context—who is gathering the information, who is analyzing it, who is disseminating it and to whom, the nature of the information, the

118. Nissenbaum, supra note 1.
119. See, e.g., GANDY, supra note 111; FERDINAND DAVID SCHOEIAN, PRIVACY AND SOCIAL FREEDOM (1992); Cohen, supra note 58; Jerry Kang, Information Privacy in Cyberspace Transactions, 50 STAN. L. REV. 1193 (1998); Reiman, supra note 100; Daniel J. Solove, Conceptualizing Privacy, 90 CAL. L. REV. 1087 (2002).
relationships among the various parties, and even larger institutional and social circumstances. It matters that the context is, say, a grocery store as opposed to, say, a job interview or a gun shop. When we evaluate sharing information with third party users of data, it is important to know something about those parties, such as their social roles, their capacity to affect the lives of data subjects, and their intentions with regard to subjects. It is important to ask whether the information practice under consideration harms subjects; interferes with their self-determination; or amplifies undesirable inequalities in status, power, and wealth.

We might agree that there is something disrespectful, even sinister, in the relentless gathering, aggregation, mining, and profiling conducted by companies like Seisint and Axciom. In other cases, contexts, or activities that are similar in form might strike most people as desirable, or at least acceptable. Consider teachers in the setting of primary and secondary education in the United States—they collect and aggregate information about students in order to assign grades. Over time, these grades are further aggregated to yield grade point averages and are combined with other information to form a student dossier, which, in some form, may be submitted to colleges or employers to which students have applied for admission or employment. A school might be judged remiss if it failed to notice that the performance of particular students had changed significantly in one way or another, if it failed to “mine” its data for other categories of change that reflected on students’ and the school’s performance.

IV. CONCLUSION

This Article develops a model of informational privacy in terms of contextual integrity, defined as compatibility with presiding norms of information appropriateness and distribution. Specifically, whether a particular action is determined a violation of privacy is a function of several variables, including the nature of the situation, or context; the nature of the information in relation to that context; the roles of agents receiving information; their relationships to information subjects; on what terms the information is shared by the subject; and the terms of further dissemination. The model is prescriptive in that it is intended to serve as a justificatory framework for prescribing specific restrictions on collection, use, and dissemination of information about people.

Although other normative theories of privacy have produced important insights into privacy and its value and foundations, they
typically are framed in overly general terms. As a result, important details that in my account give rise to systematic context-relative qualifications need to be treated as exceptions, or tradeoffs. By contrast, the possibility of context-relative variation is an integral part of contextual integrity.

By contrast, if we adopt contextual integrity as the benchmark for privacy, these context relative qualifications can be built right into the informational norms of any given context. One consequence is that privacy prescriptions, now shaped to a significant degree by local factors, are likely to vary across culture, historical period, locale, and so on. Although some might find this problematic, I consider it a virtue. As prominent contributors to the study of privacy have noted, norms of privacy in fact vary considerably from place to place, culture to culture, period to period; this theory not only incorporates this reality but systematically pinpoints the sources of variation. A second consequence is that, because questions about whether particular restrictions on flow are acceptable call for investigation into the relevant contextual details, protecting privacy will be a messy task, requiring a grasp of concepts and social institutions as well as knowledge of facts of the matter. Ideally, this approach will encourage future research into prominent and problematic domains in order to uncover how technical innovations in these domains affect informational norms.

Finally, a brief note on how to respond to violations of contextual integrity, particularly those associated with widespread adoption of technologies of public surveillance. In connection with similar questions about injustices, Michael Walzer recommends that certain types of exchanges be blocked in order to preserve complex equality. Distribution principles of one sphere should not be permitted to intrude into others, so that those who are wealthy in one sphere are not allowed to spread tyranny to others. In our own society, we experience at least some such safeguards in law and policy—such as those prohibiting monetary exchanges for various kinds of goods (e.g., votes, babies, and organs), those invalidating kinship as a basis for handing down political office, and those rejecting political office as a sound basis for favorable decisions in court; even outlawing insider trading.

120. See, e.g., Westin, supra note 98.
121. See Kang, supra note 119 (providing an exemplary naturalized analysis of a particular domain—although not couched in terms of contextual integrity).
122. See Alex Kuczynski & Andrew Ross Sorkin, For Well-Heeled, Stock Tips Are Served with the Canapés, N.Y. TIMES, July 1, 2002, at A1, B6 (“The investor Wilbur L. Ross Jr., who spends his
Privacy as Contextual Integrity

Policy and law are not the only means of preserving contextual integrity. Outside the legal arena, norms of decency, etiquette, sociability, convention, and morality frequently address appropriateness and distribution of information. Certain contexts, such as friendship and courtship, for example, as rich and important as they are, are likely to remain the purview of these non-legal systems. In certain contexts, such as that of a lawyer-client (or other professional) transaction, a middle ground has so far seemed workable—norms explicitly articulated, backed by sanctions of the relevant professional associations.\(^{123}\) When to codify contextual integrity into law, policy, and regulation is a familiar question about the scope of the law. Here, there is space to propose only that when violations of norms are widespread and systematic as in public surveillance, when strong incentives of self-interest are behind these violations, when the parties involved are of radically unequal power and wealth, then the violations take on political significance and call for political response.

weekends in the socially conscious town of Southampton, said that the people who divulge information and pass along tips are most likely concerned with improving their social status. . . . With a wink or a nod among friends and acquaintances, information heard along the boulevard is used to lubricate a promising personal or business relationship, impress a dinner table and repay a favor."

\(^{123}\) But see Jonathan D. Glater, *Lawyers Pressed to Give Up Ground on Client Secrets*, N.Y. TIMES, Aug. 11, 2003, at A1, A12 (reporting that new government rules following corporate scandals, tax evasion, and concerns over terrorism are forcing professional groups, such as the American Bar Association, to cede ground on client confidentiality). Within this approach, such a change is framed as a change in norms of distribution on the lawyer-client context.
GOVERNMENT-TO-CITIZEN ONLINE DISPUTE RESOLUTION: A PRELIMINARY INQUIRY

Anita Ramasastry*

I. BACKGROUND: ONLINE DISPUTE RESOLUTION AND COMMERCIAL DISPUTES

Lawyers, government officials, and nongovernmental organizations have become fascinated with a new type of dispute resolution in recent years. Online dispute resolution (ODR) has been the subject of a vast number of reports, scholarly articles, and newspaper articles. The term “online dispute resolution” refers to the use of information technology and telecommunications via the Internet (online technology), as applied to alternative dispute resolution (ADR). ADR, in this context, refers to dispute resolution other than litigation in courts and includes mediation, arbitration, and conciliation.

The fascination with ODR relates in part to a fascination with new forms of technology. The thought of resolving a dispute between parties in different corners of the globe using personal computers and an Internet connection is appealing. The fact that one might be able to sit at a computer in one’s pajamas and have a contract dispute mediated with an Internet merchant sounds quite attractive.

This Article first examines the use of ODR as a tool for private sector dispute resolution. It explores some of the reasons for a slower rate of uptake in business-to-consumer e-commerce disputes. The Article then suggests that a new and innovative use for ODR may be for public sector dispute resolution—between governments and citizens. The use of technology for public dispute resolution may promote access to justice in the administrative context.

ODR is simply about the use of new information management and communication tools for dispute resolution. It is essentially an offspring

* Associate Professor of Law and Director, Shidler Center for Law, Commerce & Technology, University of Washington School of Law; Vice Chair, Jurisprudence Committee, Access to Justice Technology Bill of Rights Initiative.

of ADR; ODR deploys information technology and distance communication in the context of traditional ADR processes.² Like ADR, it can provide some of the same potential advantages over litigation: greater efficiency, greater party control, and lower costs. In fact, the introduction of online technology appears to increase these advantages of ADR over litigation in terms of cost, efficiency, and convenience.³

ODR has been the subject of debate and study, almost exclusively with reference to the private sector and business disputes.⁴ E-commerce merchants, governments, and policymakers have focused on ODR as a means of resolving disputes arising between parties engaged in global e-commerce transactions. ODR has been highlighted as a means of resolving disputes in the global business-to-business market place (B2B) through the use of new online arbitration services. Similarly, for disputes arising between Internet businesses and consumers (B2C), or between two consumers (as in the context of a person-to-person transaction such

². Commentators have sometimes made a distinction between proceedings conducted exclusively online and proceedings “only” supported by different elements of online technology. In fact, there is no such clear-cut distinction. Today, all dispute resolution falls in the latter category to some extent, in the sense that online technology plays some role or other in most modern dispute resolution. Parties engaged in ADR, for example, will often use e-mail to communicate with one another or with a neutral third-party. Thus, whether a process can be called “ODR” is a matter of degree; there is a broad spectrum of ODR, with at one end proceedings using hardly any online technology and at the other end proceedings using a high degree of online technology. For this reason, it is more accurate to refer to ODR techniques or processes.


Government-to-Citizen Online Dispute Resolution

as an Internet auction), ODR has been touted as a means of quickly resolving disputes between remote parties.

Remote disputing parties can use ODR to “meet” one another online and use a web portal, e-mail, or other forms of interactive communication (e.g., web conferencing) to resolve their disputes without having to go to court and choose a forum that may be geographically inconvenient for one or both parties. Third-party neutrals (often referred to as ODR Providers)\(^5\) may or may not be involved in an ODR process. In other words, parties may resolve their disputes independent of a third-party mediator or arbitrator, or may choose to have a neutral party facilitate the proceedings.

ODR has been portrayed as particularly convenient and efficient where the parties are geographically distant because it obviates the need for traveling. In principle, ODR can be used both for disputes arising from online interactions or transactions and for disputes arising offline. However, it may be particularly apt for e-commerce disputes, where it is logical to use the same medium (the Internet) for the resolution of disputes, and where the parties are frequently located far apart. Recourse to ODR (and other forms of ADR) is also seen as a convenient way of sidestepping the complex jurisdictional issues that may arise with litigation over e-commerce disputes, particularly those involving cross-border B2C transactions.\(^6\)

The presumption that ODR is convenient and efficient pertains especially to disputes over transactions that are of the “high-volume, low-cost” type. For example, when Internet merchants are engaged in volume sales and have a certain number of disputes arising with customers, it may make sense to offer an ODR service. This can be achieved, for example, by contracting with a third-party ODR Provider that offers mediation services to resolve disputes between the merchant and its customers.\(^7\) Because online transactions between businesses and

\(^5\) The term “ODR Provider” is often used to refer to a company or entity that offers ODR services and that retains individual neutrals to engage in ODR activities.

\(^6\) This progress report adds, “alternative dispute resolution can be a practical way to provide consumers with fast, inexpensive, and effective remedies and can reduce business exposure to foreign litigation.” See Justin Kelly, White House Report Signals Consensus on ADR for E-Commerce, ADRWORLD.COM, Jan. 18, 2001, at http://www.onlineresolution.com/adrworldpress.cfm (quoting LEADERSHIP FOR THE NEW MILLENNIUM DELIVERING ON DIGITAL PROGRESS AND PROSPERITY THIRD ANNUAL REPORT, U.S. GOV’Y (2001)).

\(^7\) Nonprofit trade associations, such as the Better Business Bureau, offer such services to their members. See, e.g., BBBOnline, at http://www.bbbenline.org (last visited Jan. 12, 2004). In this
consumers are often of this sort, ODR has been highlighted as a preferred avenue for consumers who seek redress from Internet businesses with which they have dealt.8

II. THE FAILURE OF COMMERCIAL ODR IN BUSINESS-TO-CONSUMER E-COMMERCE TRANSACTIONS

Despite the novelty of ODR, has it been a success? At present, there is a lack of reliable data on the volume of disputes (arising online or offline) that have been resolved using an ODR process. Moreover, ODR Providers (which, at present, are almost exclusively private business entities) may be reluctant to disclose information about their business models and financial situation because of the number of existing competitors.

While there has been a great deal of academic and governmental interest in ODR, the emphasis on the private sector may be misplaced to some extent. To date, there appear to be few truly financially viable ODR efforts in the commercial sector.9 Some commentators have noted that ODR continues to gain acceptance, but the evidence is still anecdotal. Moreover, ODR’s acceptance may be related to its potential use for resolving offline disputes, or for providing enhanced processes for existing forms of ADR, rather than because it is being used for cross-border e-commerce disputes.10

regard, the Better Business Bureau is adapting existing offline dispute resolution services and bringing them online. New ODR Providers also provide services to Internet merchants and market places. Squaretrade.com, for example, is an ODR Provider that works with eBay, an Internet auction site, to offer mediation services for disputes arising between eBay buyers and sellers. eBay subsidizes the cost of SquareTrade, making it an affordable option for its customers.

8. The European Union and the United States committed themselves jointly to the use of ADR and ODR for e-commerce transactions in December 2000 when they issued a joint statement on this topic. The statement included the following language: “Easy access to fair and effective ADR, especially if provided online, has the potential to increase consumer confidence in cross-border electronic commerce and may reduce the need for legal action. We, accordingly, agree on the importance of promoting its development and implementation.” European Union in the U.S., Statement of the European Union and the United States on Building Consumer Confidence in E-Commerce and the Role of Alternative Dispute Resolution (Dec. 18, 2000), at http://www.eurunion.org/partner/summit/Summit0012/ECommerce.htm.

9. See American Bar Association Task Force, supra note 1, at 436–37 (“The establishment of a robust ODR industry clearly requires even greater growth and a more effective global reach, and just as clearly awaits greater financial, business, technological and legal maturity.”).

Government-to-Citizen Online Dispute Resolution

Thus, it may be premature to view ODR as a solution for consumer (or business) difficulties. The widespread use of ODR in relation to B2C transactions faces major hurdles. Some of these hurdles are primarily legal in nature. For example, European jurisdictions prevent the application of ODR (or ADR generally) if the business seeks to curtail recourse to the court system by consumers before a dispute has arisen by requiring binding online arbitration rather than permitting consumers to litigate.11

Other hurdles may be more a function of psychological, cultural, and social factors, such as wariness and a lack of “Internet literacy” on the part of many consumers.12 Additionally, the parties—particularly in conciliation and mediation processes—may dislike ODR processes because of the loss of cues and interaction available through face-to-face processes.

More generally, the problem with using ODR relates to the economics of ODR. Why is this? First, consumers may have no interest in resolving disputes over relatively small transactions. For example, if a consumer purchases a fifteen-dollar compact disc online, the cost of mediating a dispute between buyer and seller may be greater than the cost of the product itself.13

One company, Onlineresolution.com, charges a minimum of fifty dollars per hour (with a minimum of two hours) to each party for

While some ODR providers have gone out of business, other companies and projects have taken their place. For example, three years earlier there were 24 ODR companies, of which 11 had gone out of business by March 2003. In addition, most major ADR organizations, such as the American Arbitration Association and the International Chamber of Commerce, have started or are planning to start to use ODR.

Id. The report further notes that “ODR is growing in use not only because there is growth in online activities and online disputes but because ODR can also be employed for traditional offline disputes.” Id. at 185.


13. Consumers have various types of injurious experiences—some are perceived and others unperceived. More generally, only a segment of consumer injuries become what one would label complaints or disputes. Scholarly commentators often refer to a “dispute resolution pyramid” with only a segment of injuries rising from being unperceived injurious experiences to formal disputes that call for the use of ADR or litigation. See Marc Galanter, Reading the Landscape of Disputes: What We Know and Don’t Know (and Think We Know) About Our Allegedly Contentious and Litigious Society, 31 UCLA L. REV. 4, 11–36 (1983).
disputes under ten thousand dollars. After the initial two hours, the rate for each additional hour is fifty dollars. ODR Providers must also invest in a technological platform for their services, which involves an initial capital investment and possible ongoing technology licensing fees. Thus, merchants may have a disincentive (as may ODR Providers) to provide ODR services for disputes relating to low-cost consumer products or services.

To a large extent, consumers (and many businesses) want the benefits of so-called “rule of law” without having to incur the associated costs. In this context, one could define rule of law as the concern for ensuring that the dispute resolution mechanism complies with principles of due process; i.e., that decision-making is non-arbitrary, non-capricious, predictable, and transparent. Thus, even with ODR, concerns of due process remain prominent.

Consumers, however, may want due process at an affordable price. This has lead organizations such as the American Bar Association to recommend that merchants who wish to enhance consumer confidence in e-commerce transactions improve complaint handling procedures rather than investing in ODR processes.

A related issue is funding: consumers may want low-cost schemes but the ODR Provider needs to cover its costs and deter frivolous complaints. Who should pay arbitrators or mediators in cases involving consumer small claims? Consumers may be unable or unwilling to foot their bill or even a substantial part of it—especially when the dispute is over a small purchase or the claim is of a low value.

Funding appears to be the crux of the problem. There are several ways that ODR can be funded, including user fees (bilateral or...


Disputes under $10,000 (and non-monetary disputes desiring Experienced Neutral services) are handled by the Online Resolution’s Experienced Neutrals Panel. To begin, complete a Dispute Information Form. Online Resolution will then seek to contact the other party. If the other party is agreeable to participating, Online Resolution will charge by Visa or MasterCard a $50/hour fee for each participant for up to two (2) hours of dispute resolution services. Total: $100 per party. Beyond this 2 hour minimum, fees are $50/hour from each party. If the other party does not agree to the selected process, there is no charge and your information is deleted from our system.


17. American Bar Association Task Force, supra note 1, at 442–43.
Government-to-Citizen Online Dispute Resolution

unilateral), membership fees, and external sources. The bilateral fee model involves both parties to a dispute (e.g., merchant and consumer) paying for the ODR services. The problem with the bilateral user fee model is that for small- or medium-sized claims, the costs may be too high for the claimants (especially if the claimant is a consumer).

By contrast, if the fee model is unilateral (one party agrees to underwrite the cost of the dispute resolution), or based on membership, only one party pays. In the B2C context, the merchant is the paying party. This raises questions about the independence of the ODR Provider. If an ODR Provider is being retained or hired by the merchant, questions may be raised about the independence of the neutral parties—because they are dependent, to some extent, on the merchant for their fees and livelihood.

Of course, there are ways to ameliorate the independence problems raised by the merchant paying for the ODR process. For example, numerous small claims of the same kind could be merged into one relatively large claim for treatment by the ODR Provider. However, some ameliorative strategies will raise further problems. For instance, were an ODR Provider to meet its costs through substantial sponsorship from business, its ability to act independently and impartially—and just as importantly, its ability to be seen as acting independently and impartially—might well be compromised.

The unilateral fee model, however, is used in many contexts. Many merchants belong to merchant trust mark programs, such as the Better Business Bureau, that offer arbitration services to members in disputes arising with customers. Merchants display a certain membership seal in their physical store, or a web seal or trust mark online, as a sign that they


19. In this context, dispute resolution is a service offered by one party to its clients or customers. The dispute resolution service is made available on the basis of membership in a trade or professional association.


The trouble with this model is that it raises problems of independence. If one party pays exclusively or much more than the other party for a dispute resolution, a real or at least a perceived bias inevitably appeals. It is a form of business affiliation that should be avoided, because it lessens trust, and maybe also the quality of justice.

Id. See also Lucille M. Ponte, Throwing Bad Money After Bad: Can Online Dispute Resolution (ODR) Really Deliver the Goods for the Unhappy Internet Shopper?, 3 Tul. J. TECH. & INTELL. PROP. 55, 67 (2001).
adhere to a code of business practices. These practices include participation in ADR or ODR if a dispute arises.

SquareTrade, another ODR Provider, has partnered with the Internet auction site, eBay, to provide negotiation and mediation services for buyers and sellers involved with eBay auctions. The cost to a consumer or user of the auction site is only fifteen dollars. eBay underwrites this program and pays SquareTrade the rest of its fees.21 Similarly, many companies that operate offline require consumers to participate in mandatory arbitration proceedings in the event that a contractual dispute arises. In such circumstances, the merchant may pay the arbitration fees for the proceeding.

Another potential issue is a tension between transparency and confidentiality. Prospective parties to B2C transactions will tend to want information from ODR Providers on how previous disputes have been handled, including the outcomes and reasoning applied. Transparency at this level will help meet the general need for prescriptive guidance. Yet, actual parties to disputes will frequently want the nature and outcomes of the ODR proceedings kept confidential. Encouraging transparency between the parties will buttress the integrity of the proceedings.

Businesses are often reluctant to participate in ODR as well. In some circumstances, businesses may ignore online complaints. One reason why they can do this is because there is no clearly established online customer community that can collectively react to questionable business practices.22 At present, businesses may find that providing a strong in-house complaints handling system may be more cost effective than

21. SquareTrade offers additional ODR services outside of the eBay context. These cost more money for end users. It offers online mediation services for buyers and sellers of real estate, for example. The cost for such a service is $175.00 plus mediation fees of $100 per hour per party. Such mediations typically last three to six hours. See SquareTrade.com, Help—Dispute Resolution, at http://www.squaretrade.com/ent/jsp/hlp/re_fees.jsp;jsessionid=3xpaxuxel1?vhostid=chipotle&stmp =car&cntid=3xpaxuxel1#cost (last visited Jan. 12, 2004). While this may be cost effective for real estate disputes involving property valued in the thousands of dollars, such a model is less attractive for B2C e-commerce disputes.

22. See Rufus Pilcher, Trust and Reliance—Enforcement and Compliance: Enhancing Consumer Confidence in the Electronic Marketplace 115 (2000) (unpublished J.S.M. thesis, Stanford University), at http://www.oecd.org/dataoecd/0/18/1879122.pdf (discussing collective punishment by a seller’s customer community and its absence in e-commerce: When not only the defrauded individual consumer refrains from repeat transactions with a merchant, but all or at least a large group of consumers boycott that merchant, the threat of a sanction will be considerably more powerful. As a consequence of such coordinated approach, a merchant who cheats one consumer risks to deprive it from future revenues of potential dealings with all consumers. This mechanism is known as collective punishment.)
Government-to-Citizen Online Dispute Resolution

resorting to an outside ODR Provider.\(^{23}\) Additionally, studies have revealed that Internet merchants are not necessarily conforming to best practices with respect to handling consumer e-commerce complaints.\(^{24}\)

Finally, less costly alternatives to ODR exist for consumers including (at least in the United States) the use of credit card charge-back mechanisms and Internet escrow services as a means of securing payment and performance of e-commerce contracts.\(^{25}\) Such mechanisms may also be more effective for ensuring that merchants comply with various outcomes rather than relying on the merchant to voluntarily comply with a mediated settlement. With a credit card, for example, if a U.S. consumer/cardholder has a legitimate contractual dispute with the merchant, the loss will be passed back to the merchant.

In the future, ODR techniques may become more common in the e-commerce context. As ODR Providers offer a range of services for online and offline disputes, the cost of providing B2C ODR may decrease as the overall market for ODR grows. Moreover, ODR may be appropriate for some types of consumer claims, such as insurance claims, where automated settlement software has been used to create settlements between insurance companies and consumers.\(^{26}\) At present, however, it is difficult to predict the future of B2C ODR, especially given the lack of statistical and financial information that is publicly available. Rather than B2C, the demand for ODR in the private sector may emerge for B2B disputes generated offline and online.\(^{27}\) Businesses are able to fund the cost of ODR bilaterally, thereby eliminating the funding constraints discussed above.


Internet merchants are on the front line of dealing with consumer e-commerce complaints. Realistically, consumers often turn to a merchant and its in-house customer service department or complaint handling process first if a problem arises from an e-commerce transaction. The Internet merchant will have familiarity, data, and expertise that will allow it to provide effective complaints resolution processes for the consumer. As for cost, a strong and professionally conducted in-house complaint resolution program is inherently less costly than the use of third-party ODR or ADR service as a means of resolving consumer complaints.

Id.

24. Id. at 431.

25. Id. at 457.


27. Id. (discussing role of ODR Providers for larger value commercial disputes).
III. THE PROMISE OF ODR: PUBLIC SECTOR DISPUTE RESOLUTION

Could technology-assisted dispute resolution be useful in other contexts, outside of the e-commerce arena? The answer is yes—with respect to resolving government-to-citizen (G2C) disputes. To date, there has been little focus on the promise of ODR in the public sector.

Many reasons for the lack of ODR deployment in the private sector may be eliminated when ODR is deployed in the public sector.\(^\text{28}\) When ODR is employed as a public good rather than as a private commodity, the economics of running a successful business become less of a concern.\(^\text{29}\) The government could assess the cost of deploying technology in order to offer dispute resolution mechanisms for its citizens and decide whether to offer such a service in light of competing priorities within the justice system.\(^\text{30}\)

What would make ODR attractive to government entities? At present, it may be premature to advocate full-blown court-based adjudication that is conducted solely in cyberspace. Although there have been some moves to create online courts (at least in the civil context),\(^\text{31}\) there are much greater considerations at stake with respect to transferring litigation proceedings, which involve complex issues relating to documentary evidence, witnesses, and the role of counsel, for example, into the online context.

Where might ODR play a role in government-based adjudication? At present, a starting point might be situations in which the government is

\(^{28}\) See Schultz, supra note 18, at 7 (arguing that ODR in the public sector creates greater accountability since judges are public servants).

\(^{29}\) Id. at 9.

\(^{30}\) See UNCTAD, supra note 10, at 191.

A major function of government agencies is the resolution of disputes between citizens and government, or between citizens and other citizens. In addition, many government functions such as rule making, may involve trying to achieve consensus among interested parties, a very familiar dispute resolution goal. While ODR has in the past few years been concerned mostly with the private sector, increasing efforts in the areas of e-government and e-democracy are focusing attention on the value of ODR.

Government-to-Citizen Online Dispute Resolution

already engaged in high-volume administrative proceedings that are currently adjudicated remotely. If the state has already provided a mechanism for citizens to appeal or contest government actions remotely, by mail for example, it may streamline the process by offering similar functionality with the assistance of technology.

Citizens routinely have administrative appeals and other types of disputes with government entities at the local, municipal, state, and federal levels. Disputes arise in multiple contexts. For example, citizens appeal parking tickets or participate in mitigation hearings for such citations, as well as for tax assessments, zoning permits, etc.

In many instances, governments have an administrative system in place to deal with citizen appeals and to adjudicate them either by mail or in a person-to-person context.32 Citizens can, for example, often appeal the issuance of a parking ticket or a moving violation in a court or administrative hearing. Alternatively, they can often choose to have a hearing by mail.33

32. As an Australian government report notes:
   For non-consumer disputes, the role played by government in online ADR should be driven by access to justice and quality of service concerns. Where governments already provide face-to-face or telephone ADR services, there is a strong argument for extending these services online: both to meet public expectations and to improve access to services for a number of groups, including rural and regional, hearing impaired and other disadvantaged citizens.


Pierce County, Washington has also lauded the use of hearings by mail:
   Rather than require a defendant to personally appear in a courtroom to present testimony for a traffic violation which typically takes five minutes, the Pierce County District Court No. One has provided a process for the defendant to submit a written statement for consideration.
   Washington State court rules require an infraction hearing to be scheduled within 120 days of the violation date. Early in 2000, it became apparent to the court that the time frame was unlikely to be met. Available hearing dates were consistently set at 15 weeks, leaving little flexibility for the court or the defendant. Requests for hearings had also increased. Between 1995 and 2000, the percentage of cases that required a hearing increased 6 percent overall (25 percent in 1995 to 31 percent in 2000).
   Without a remedy, the court faced the possibility of dismissing cases for failure to schedule the hearing within the mandated time frame.
   Simply adding more court room time to accommodate the increased need for hearings was not feasible. The court had experienced a reduction in judicial positions in 1999, and an increase in 2000 was not an option.
By expanding the methods by which citizens can have claims involving the government adjudicated through the use of ODR, one may ultimately expand access to justice by creating greater opportunities for citizens to interact with the state and to have their grievances resolved in a timely fashion. Similarly, this may allow for the state to deal with a larger number of cases because technology can be deployed to speed up a traditionally paper-based process.

Another instance where ODR may be useful is for citizens who have problems arising from online transactions that they conduct with government entities. Many government services, such as renewing car licenses, requesting vital records, and paying taxes, are available online. Errors are bound to arise with such transactions. An online or “electronic” ombudsman could serve as an intermediary for resolving problems that arise with G2C electronic transactions.34

Have governments deployed ODR for public sector dispute resolution? To date, the general answer appears to be no. In the United States, the Internal Revenue Service recently announced that it would launch a new service to resolve tax disputes online. If this occurs, it would be the first such federal government initiative in the United States.35 At the municipal level, New York City offers “Parking

To address the expanding quagmire, the Pierce County court turned to a Washington State court rule that allows individual courts to conduct certain hearings-by-mail. A local rule was adopted authorizing infraction hearings based on written statements given under penalty of perjury. The information sheet for infractions was revised to include the hearings-by-mail process. Additional forms were developed to standardize the process, provide detailed information to the defendants, and lessen the confusion for defendants and staff. Defendants typically requested hearings for the infractions through the mail or at the public counters. Since providing consistent information was important to minimize staff workload and maximize the new procedure’s impact, counter clerks were trained to explain this option to defendants requesting a hearing and provide forms as appropriate.

The program has been extremely successful. The timeframe for scheduling hearings dropped from 15 weeks to 12 weeks; of the 10,917 infraction hearings conducted by the court during the first 11 months of 2001, 2,705 were hearings-by-mail, representing 25 percent of the hearings; mitigation hearings conducted by mail represented 35 percent of all mitigation hearings and 19 percent of all infraction hearings; 77 percent of the defendants receiving information with their hearing notice responded with a written statement; payments are received sooner with the hearings-by-mail cases due to a quicker response by the court; and possible case appeals are reduced, as they are not allowed (under state law) from a decision reached through the hearings-by-mail process.


35. David Cay Johnston, IRS Set To Resolve Disputes Online, N.Y. TIMES, Dec. 1, 2003, at C7 (abstract available at 2003 WL 67281935). Although the headline characterizes the IRS service as a
Government-to-Citizen Online Dispute Resolution

Violating Hearings by Web” as a counterpart to its hearings by mail program for parking infractions.36

In developing or transition economies, governments are exploring the use of ODR for public dispute resolution. In Hungary, for example, the federal telecommunications ministry may engage in an ODR pilot program to resolve consumer complaints with respect to telecommunications service providers.

The European Union is also exploring ways to facilitate public sector dispute resolution. The Information Society Directorate within the European Commission is exploring the potential use of ODR in the e-government sector for disputes that arise relating to data protection, copyright infringements, and domain name.37 It has established a European Extra-Judicial Network (EEJ-NET) that links together many dispute resolution services offered by member states within the European Union. European countries have a tradition of publicly-funded and privately-operated ombudsman and other ADR providers who resolve consumer disputes with businesses in the private sector.38 Many of these existing offline dispute resolution services are beginning to consider offering ODR components. EEJ-NET, as a government clearinghouse, could help to facilitate consumers’ ability to access ODR services.

The European Union has also funded a pilot program for private sector dispute resolution known as ECODIR (Electronic Consumer Dispute Resolution). ECODIR, which is currently offered free of charge, allows businesses and consumers to use this ODR service to resolve e-

new form of online dispute resolution, it appears that IRS may only provide tax practitioners access to their records online. This will assist in clearing up errors or disputes, but is not full scale ODR.

36. The New York City Hearings by Web web site indicates that not all claims may be resolved via the Internet and e-mail: 
   Using this system is only a request to the Department of Finance for a Hearing. You will not receive an immediate determination. Your case will be reviewed by an ALJ. If we are able to adjudicate your claim based solely on your e-mail, you will receive an e-mail confirmation and an ALJ's written decision by conventional mail. If we are unable to adjudicate your claim, you may be asked to either provide additional information / evidence to support your defense or have a by mail or in-person PVO hearing as described above.


commerce disputes. While this is not G2C ODR, it provides a model for government-subsidized and implemented ODR.

The Singapore court system offers an ODR service referred to as e@dr. This service provides online mediation and arbitration for parties to an e-commerce transaction. As with ECODIR, the e@dr service is not available for G2C disputes.

Governments have made some very preliminary forays into offering ODR processes for citizens. At present, however, the examples are few.

IV. G2C ODR AND ACCESS TO JUSTICE CONSIDERATIONS

One of the main focuses when looking at access to justice and technology is to examine the types of commitments that a government should make when it deploys technology as part of the legal system. If deployed with proper safeguards, ODR does facilitate access to justice.

How would one ensure due process within the context of G2C ODR? First, such ODR would be regulated—not unregulated. Thus, just as current hearings by mail are regulated through statutes and rules, ODR would be likewise regulated. Additionally, to the extent that administrative procedures are similar to state or federal administrative procedure laws, there would be additional safeguards in place with respect to a citizen’s right to be heard.

In addition, in devising new public sector ODR, one could adapt principles that have been developed in the private sector ODR context and apply them to the public sector process. There are a series of guidelines and principles for B2C ODR that government bodies, nongovernmental organizations, and industry have developed as a way of articulating best practices for private ODR. In many senses, these principles are meant to embody principles of access to justice, fairness, and due process that replicate principles derived from the traditional legal system. Thus, they embody principles that are highly relevant for public sector ODR.


41. See Schultz, supra note 18, at 10 (“An important reason why we need ODR is that it constitutes an access to justice, sometimes the only access to justice that is reasonably available to the parties. . . . ODR may be more than just a tool that solves disputes. It may also be a way to provide justice.”).
Government-to-Citizen Online Dispute Resolution

Some of the principles that appear in statements of ODR best practices include:

- **Transparency**—ODR programs should provide readily-accessible information about all aspects of their services;
- **Independence**—ODR programs should operate independently of business interests;
- **Impartiality**—ODR programs should operate without bias favoring business interests;
- **Effectiveness**—there should be mechanisms to ensure business compliance with ODR outcomes;
- **Fairness and Integrity**—ODR programs should observe due process standards ensuring that each party to a dispute has equal opportunity to express its point of view;
- **Accessibility**—ODR programs should facilitate easy use by consumers;
- **Flexibility**—ODR programs should permit adaptation of their procedures to suit the circumstances of the particular dispute at hand; recourse to courts by consumers should not be precluded unless by prior and equitable agreement; and
- **Affordability**—ODR programs should be affordable for citizens or consumers, particularly in light of the amount of compensation being sought.

For the principles listed above, one could substitute “government” for “business” to achieve a similar list of principles for G2C ODR. Given the “legal” and possibly binding nature of G2C ODR, self-regulatory principles or best practices would not be enough. One would still need formal rules and procedures.

In a broader sense, G2C ODR would involve public magistrates or officials, it has the potential to instill greater confidence and trust in the proceedings than ODR in the private sector, where the problem of unilateral fees being paid by the industry creates perceptions of bias or partiality in the process. To the extent that citizens feel confident in remote adjudicatory processes to date, ODR may enhance trust and confidence by allowing citizens to interact more directly with the adjudicator and the state—using e-mail or online real time communication.

Of course, the concept of ODR raises other concerns that must be considered. First, not everyone will have the means to access technology
to engage in ODR processes in the administrative context. Creative use of public sector kiosks and terminals, perhaps even in government buildings or courthouses, might ensure greater access to ODR than would be feasible if individuals were required to deploy their own technology as a precursor to benefiting from ODR. Second, as with private sector ODR, issues of computer literacy, culture, language, and other factors will certainly come into play with respect to ODR in the public sector. This suggests at least initially that ODR cannot supplant in-person hearings or remote hearings by mail; ODR becomes an alternative or supplement rather than an end in itself.

In closing, one hopes that governments will explore the potential for ODR as a public good rather than cordon its possible use in the private sector. When one sees dispute resolution as serving an access to justice purpose rather than merely as a private complaint resolution mechanism, the potential for ODR becomes much greater.
RISE OF THE MACHINES: JUSTICE INFORMATION SYSTEMS AND THE QUESTION OF PUBLIC ACCESS TO COURT RECORDS OVER THE INTERNET

Gregory M. Silverman*

The machines are coming. They have been slowly taking up positions in our courthouses for more than a quarter of a century. With each passing year, they are becoming faster and more powerful. They are evolving intelligence and the ability to communicate with each other. With their assistance, the justice system is becoming more efficient—an integrated network. Soon the courts, justice agencies, law enforcement, correctional facilities, social services, and treatment providers will be able to interoperate seamlessly. Moreover, the justice machines will be able to reach out and assimilate into their network the millions of machines connected to the Internet and owned by the public, enabling the exchange of information on a scale and with an ease never before imagined. With the assistance of the machines, the myriad and diverse members of the justice and public safety communities together with the public will evolve into a single complex whole that could dedicate itself to creating a more humane and just society comprised of better informed individuals to whom they are genuinely accountable.

Some, however, are not so sanguine. They view the rise of the machines as ominous and foreboding—threatening our privacy and the erosion of human freedom and autonomy. As information flows through an integrated justice system out to the public, they worry that human actions and relationships will be subverted by these machine connections: that people will change their behavior out of fear that their frailties, misfortunes, and “unusual” preferences will be revealed and reviled, exposed and ridiculed. To avoid this travesty, they argue, privacy must supervene our traditional commitment to public access: we must tolerate some opacity in our governing institutions and limitations

* Associate Professor, Seattle University School of Law. M.A. 1984, J.D. 1987, M.Phil. 1991, Columbia University. For helpful comments on earlier drafts of this Article, the author gratefully thanks Keith Aoki, Pat Brown, Steve Burnett, Vince Chiapetta, Mark Chinen, Maggie Chon, Annette Clark, Anne Enquist, Bob Gomulkiewicz, Lilly Kahng, Jack Kirkwood, Evan Lenz, Lydia Loren, Bob Menanteaux, Joe Miller, Chris Rideout, Veronique Silverman, John Strait, Jane Winn, and Peter Winn. I would also like to thank my research assistants Yvonne Mattson and Keith McGahan for their excellent assistance on this project.
on our access to them. According to these doomsayers, inexpensive and convenient public access to court records over the Internet must be abjured if we are to preserve what remains of the collapsing catacombs of personal privacy beneath an increasingly mechanized and hostile world.

Fortunately, the dilemma between privacy and public access is a false one, nor the world as Orwellian as the first paragraph might suggest. The same technology that heralds unprecedented public access at minimum cost and maximum ease also enables an automated intelligence that is capable of understanding and processing data in sophisticated and nuanced ways yet to be generally appreciated outside the circle of technologists who currently work with it. This technology—Extensible Markup Language (XML) and its family of related browsers, parsers, processors, and standards—permits information in court records to be shared with the public at the courthouse and over the Internet while respecting the legitimate privacy interests of litigants and others who come before our courts. With it, humans really can achieve a more just and humane society—one in which they remain clearly in control.

In the first half of the present Article, I introduce the reader to this technology (Part II) and its likely role in evolving justice information systems (Parts I and III). In the second half of the Article, I enter the debate over whether the public should be permitted access to court records over the Internet. After explaining the origins, history, and principal sides of this debate (Part IV), I argue, first, that when used properly, XML permits the public to have access to court records over the Internet while promoting public safety and protecting personal security (Part V) and, second, that the presence of discrediting and embarrassing facts in a case file does not justify limiting public access to court records over the Internet while permitting unlimited public access at the courthouse (Part VI).

I. CASE MANAGEMENT INFORMATION SYSTEMS: FROM INDEX CARDS TO XML

Before the development of electronic databases, tools for managing large quantities of information included paper notebooks, ledgers, and variously sized wooden cabinets and cardboard boxes filled with paper index cards. During this period, if one wanted to determine if a person or

---

1. See infra Part VII for a more detailed overview of my argument.
Rise of the Machines

business were involved in litigation, one had to travel to the local courthouse of a particular jurisdiction and scan the columns in a court ledger or flip through a narrow drawer of carefully alphabetized index cards. Using paper-based information management tools such as these, court personnel developed often elaborate and ingenious systems for recording and tracking information essential to the daily operation of the court.

With the rise of modern computing and electronic databases, courts began to migrate from paper-based systems to electronic information systems. Frequently referred to as case management information systems, these systems are in fact database management systems (DBMSs) that court personnel use to input, store, manipulate, display, and print information relevant to such daily court tasks as case filing, calendaring, docketing, case maintenance, recording judgments and sentences, accounts receivable and collections, cashiering and creating receipts, trust accounting, checking and banking, failure to appear and warrant processing, as well as management and statistical reporting. As one can see from the preceding list, the functionality and use of these systems extend far beyond case management narrowly construed. Accordingly, such systems might better be called court record management systems or simply judicial information systems (JISs). Both labels are used in discussions of court automation and integration.

Of the more than 16,000 courts in the United States, not all immediately migrated to electronic case management information systems. Initially, few courts had both the funding and the technical expertise to make this transition. The early case management information systems required large and expensive mainframe computers running proprietary, often custom-written, software that was hard to modify, and could only be accessed through dumb CRT (cathode-ray tube) green-screen terminals. With the development of microcomputers in the late 1970s, these early information systems running on mainframes were either replaced or supplemented with personal computers on the desktop of select court personnel. As the price of computer hardware and storage fell and an understanding of the new technologies increased, more and more courts forswore their paper-based systems and entered the information age. While these personal desktop computers ran “productivity” software that permitted the

recording and management of information, the data entered into these computers became “trapped in the machine” and could only be shared with others through printed lists and paper reports. With the introduction of modern networking technologies such as Ethernet in the mid to late 1980s, a court’s isolated desktop computers could be connected together into a courthouse local area network (LAN). Using software that causes a personal computer to emulate a dumb CRT terminal, even a court’s existing mainframes could be tied into this courthouse network—with some effort. By the late 1980s and early 1990s, most courts—to the extent that they could afford it and the courthouse facilities permitted it—were stringing their desktop computers together with network access cards and Ethernet wiring.

Without a doubt, connecting a court’s computers into a single LAN was a real achievement and an important milestone in the history of court automation and integration. For the first time, court personnel could access, in real time and from their own desktop computers, the various electronic databases running on different computers within the courthouse. Nonetheless, network connectivity by itself could not provide a single integrated view of all the information and data relevant to a particular case or even of a particular event within a case. Courts, having automated their operations while the computer and software industries were still maturing, had different court processes and functions that were automated at different times, on different platforms, by different software programs. As a result, in many courts, the data relevant to any particular case was not found conveniently centralized in a single database management system but in several such systems and software programs distributed over the entire network. For example, to access the schedule for a particular case, one might have to consult a stand-alone calendaring program; to check whether a party had filed a document in that same case, a stand-alone docketing program; and to confirm payment of a court fee, a stand-alone accounting program. Before one could achieve a single, integrated view of all the information and data relevant to a case, one would have to not only network all of the

3. While one could transfer the data file from one desktop computer to another one using a floppy disk, this practice was not advisable as it posed significant risks to the accuracy and reliability of the information system. The transferred copy, for example, would be immediately out-of-date (since it could not be updated in real time as new information was added to the original), and created the possibility of divergent and conflicting data stores (if one added new information to the copy rather than the original).
computers storing such information but integrate the programs and information systems running on these machines as well.

While technically daunting, the benefits that result from programmatically integrating a court’s information systems are clear: cost savings, error reduction, and improved performance. Simply integrating the diverse information systems in a single courthouse reduces a court’s operating costs significantly. According to one study, approximately fifty percent of a court’s operating expenses may be attributed to the handling and storage of paper documents.\(^4\) While the introduction of electronic information systems reduces the volume of paper that court personnel must handle, integrating these information systems reduces this volume even more. The ability to enter data and access electronic records over a single court network obviates the need for paper forms as well as printed lists and reports that are used to “paper over” or “bridge” the information chasm that separates these islands of data.

In addition, integrating a court’s information systems reduces the cost of maintaining duplicative records across multiple systems. When a court’s information systems are not integrated, each system must contain information that duplicates records residing on other court systems. For example, the docketing database, electronic case index, and court accounting program—to name just three applications—will all need to contain information about parties and their attorneys. When this information changes, it must be updated separately on each system. If, however, these three systems could share data with one another, such duplicative recordkeeping would be unnecessary; each program could simply call up the information from a central database as needed. Court operating expenses would be reduced to the extent that court personnel would no longer have to repeatedly enter the same information into different information systems. Moreover, when information changes, all systems could be kept current by simply updating a single record in the central database.

Integrating court information systems in a manner that permits electronic filing also lessens personnel costs by reducing the number of court personnel needed to cover the counter in the court clerk’s office. Finally, the ability to route electronic documents to judges, parties,

---

attorneys, law enforcement officers, witnesses, and others reduces postage and delivery expenses.

Besides these financial savings, integrating a court’s information systems also reduces data entry errors and related problems. When data must be entered more than once, there is a greater likelihood of mistakes: information may be entered inaccurately or incompletely. If such information were stored as a single record on a central database, redundant records would not be needed and the opportunity for data entry errors obviated. Integrating a court’s information systems also improves the court’s ability to meet the American Bar Association Standards Relating to Court Organization and adjudicate the cases before it “justly, promptly, effectively, and efficiently.”

Judges and other court personnel can immediately access the information they need when they need it. Delays due to lost or incomplete paper files are avoided. Court scheduling conflicts are identified and prevented. Electronic reminders of required tasks are displayed. The gains in productivity and quality of service that result from a court’s integrating its information systems are as varied as they are numerous.

Even more benefits accrue as courts integrate their information systems with the information systems of other courts, justice and public safety agencies, social services, and treatment providers. An especially compelling example of the benefits to be achieved by integrating court case management information systems with information systems run by a sheriff’s department or department of corrections is provided by the Los Angeles Sheriff’s Department: an inmate serving time on misdemeanor traffic violations was mistakenly held for nine months after he should have been released. The inmate, thirty-three year old Thao Quoc Huynh, was sentenced to serve four concurrent 150-day sentences for driving under the influence, hit-and-run, and two other traffic violations. However, due to confusion over paperwork received from the court, the inmate ended up serving four consecutive sentences, keeping Huynh behind bars for 271 days beyond his correct release date. Sheriff Sherman Block explained this error by noting that the paperwork that was sent to the jail from the courthouse was ambiguous,
indicating that Huynh was to serve both concurrent and consecutive sentences. Unfortunately, the clerk at the jail failed to contact the court for clarification regarding Huynh’s sentences and simply processed the inmate as if he had been sentenced to serve the sentences consecutively.8

Additional research by a reporter for the L.A. Times revealed that the problem of over-detaining inmates has been going on for years. In 1997, nearly 700 inmates were held in county jails for an average of 6.9 days past their ordered release dates. One inmate was held 260 days too long; two others were held for 90 days or longer. All too aware of the county’s financial liability on the issue, the department’s risk management unit in 1997 paid nearly $200,000 to 548 inmates who were incarcerated for a total of 3,694 days beyond their sentences—on the condition that they agree in writing not to sue.9

To address this problem, Sheriff Block stated that his “department is establishing a computer system that will link the Inmate Reception Center with courthouses, eliminating the need to manually process thousands of pieces of paperwork at the jails each night.”10

In addition to reducing the number of human custodial errors by providing the relevant justice agencies with accurate information regarding court judgments and sentences, integrating court and justice agency information systems provides the courts with better and more timely information concerning parties before it, thereby enhancing the quality of the courts’ decision-making. Integrating the courts’ and justice agencies’ information systems also facilitates court compliance with various reporting, record checking, or data collection requirements imposed by state and federal legislation. For example, on the federal level, the National Child Protection Act of 1993,11 the Brady Handgun Violence Prevention Act,12 the Lautenberg Amendment13 to the 1968 Gun Control Act,14 the Jacob Wetterling Act (including Megan’s Law),15

8. Id.
9. Id.
10. Id.
13. 1997 Omnibus Consolidated Appropriations Act, id. § 922(g).
the Pam Lychner Act,\textsuperscript{16} and the 1994 Violent Crime Control and Law Enforcement Act\textsuperscript{17} all impose reporting, record checking, or data collection requirements on state courts or justice agencies. Compliance with these requirements necessitates close coordination and data sharing between the courts and justice agencies. Similar coordination and data sharing are also required to comply with state laws imposing enhanced sentencing for repeat offenders (so-called “three strikes” legislation) as well as laws enacted pursuant to Megan’s Law requiring registration of sex offenders and some form of community notification upon their release.\textsuperscript{18} Finally, to the extent that such reporting and data sharing prevent dangerous criminals from being released due to inaccurate or incomplete information, halt the sale of guns to convicted felons, and apprise communities when known sex offenders move into their area, integrating the information systems of the courts and justice agencies also improves public safety.

Significant benefits also accrue when courts integrate their information systems with social services, such as welfare and child support services, and treatment providers. To a greater and greater degree, courts are forming an essential and central hub for the delivery of social services. With the rise of alternative sentencing, domestic violence courts, drug courts, and a variety of diversion programs, courts find themselves coordinating closely with traditional social services and treatment providers. When a court places a defendant in a drug rehabilitation or anger management program, it must periodically hold review hearings and monitor the defendant’s attendance and progress. In a drug program, the results of a defendant’s regular drug tests must also be collected. Alternative sentencing, such as community service, requires a court to remain in regular contact with the organization or agency overseeing the defendant’s work. Assigning defendants to a residential treatment center requires the court to monitor the availability of beds at that center. The need to collect and manage new forms of information in an increasingly large number of cases places an enormous burden on court personnel and the limited resources at their disposal. These burdens may be substantially reduced and the evolving social services functions of the court greatly expedited if courts are able to

\begin{itemize}
\item[\textsuperscript{16}] Pam Lychner Sexual Offender Tracking and Identification Act of 1996, \textit{id.} \textsection{} 14072.
\item[\textsuperscript{17}] National Stalker and Domestic Violence Reduction, 28 U.S.C. \textsection{} 534.
\item[\textsuperscript{18}] See, \textit{e.g.}, CAL. PENAL CODE \textsection{} 11105 (West 2003); WASH. ADMIN. CODE §§ 446-20-500 to -510 (2003).
\end{itemize}
share and exchange data with the social services agencies and treatment providers that support the justice and public safety communities.

In light of the myriad and substantial benefits that accrue when isolated judicial information systems are integrated within and between courts, justice and public safety agencies, social services, and treatment providers, the justice and public safety communities should clearly be striving to achieve as complete an integration of judicial information systems as possible. Nor has this conclusion been lost on national and state policymakers. In 1999, the National Task Force on Court Automation and Integration declared that “[t]he time has come to improve the quality of the nation’s justice system by improving information exchange within the system.”

Making a similar point, the Judicial Information System Committee (JISC) for the State of Washington has observed that the software applications composing the State’s current JIS Application Portfolio “are, to a significant degree, islands of automation” with “significant levels of functionality [that] remain redundant and isolated from the rest of the court enterprise.” The JISC is working to replace these islands of automation with “a strategic, enterprise-wide court information system in the state of Washington” that will “automate and support the daily operations of the courts” as well as “maintain a statewide network connecting the courts and partner criminal justice agencies to the JIS database.” In order to achieve such an integrated, statewide judicial information system, the JISC concluded that the courts and justice agencies must migrate from existing legacy systems to applications, using an object-oriented Web-based architecture.

19. Nat’l Task Force on Court Automation & Integration, supra note 4, at 2. The National Task Force on Court Automation and Integration was assembled to guide the Court Information Systems Technical Assistance Project, a joint project of the Bureau of Justice Assistance (BJA), U.S. Department of Justice (DOJ); SEARCH, the National Consortium for Justice Information and Statistics; the National Center for State Courts (NCSC); the National Association for Court Management (NACM); and the Conference of State Court Administrators (COSCA).


21. Id.


23. This migration is necessary because “[t]he legacy applications are complex, difficult to use, inflexible, and do not adequately support changing business needs such as electronic filing and data
Washington has repeatedly won national awards for its leadership in e-government and public sector use of technology; if it is dissatisfied with the current level of JIS integration and data exchange within and between courts, justice and public safety agencies, social services, and treatment providers, then this suggests an endemic problem among courts across the nation.

Nor are the reasons for this state of affairs difficult to discern. The obstacles to successfully creating a statewide judicial information system that can share and exchange data with all the constituencies that a court serves are significant. To begin with, each state must develop its own judicial information system. A state cannot simply borrow a judicial information system developed by another state because each state’s governmental structure, as well as the set of legal, political, institutional, and technical requirements that a judicial information system must satisfy, are unique. 24 Nor should one underestimate the complexity of planning for such a technology project. Justice system integration requires significant levels of coordination across legislative, executive, and judicial branch agency lines, as well as across federal, state, and local jurisdictional boundaries. 25

Another obstacle arises from the manner in which most state courts govern themselves. Unlike the hierarchical structures found in executive branch agencies and private corporations, the governance structure of state courts is typically more fragmentary and diffuse. As a result, for a new initiative or project to succeed, individual judges must step forward and assume the mantle of leadership. With rising caseloads and the other obligations that accompany elevation to the bench, judges have scant time to take responsibility and ownership of complicated, multi-year technology projects. Moreover, even if judges and other court personnel were willing to spearhead the development of an integrated, statewide judicial information system, in a time of tight state budgets the funding for such an expensive undertaking would be at best precarious. Even in

---

24. As the National Task Force on Court Automation and Integration has recognized, “[s]tates are frequently dissimilar in the structure of the judicial branches and jurisdictions assigned to their courts, which makes it difficult to develop transferable automation solutions and create national data standards that are relevant from state to state.” NAT’L TASK FORCE ON COURT AUTOMATION & INTEGRATION, supra note 4, at 7.

25. Id. at 37.
states that can afford such an initiative, court funding is usually a relatively low priority, and courts are forced to compete with executive agencies and popular legislative programs for the limited discretionary spending included in a state’s annual budget. It is difficult to develop the necessary institutional commitment and expertise if continued funding for the project is considered uncertain.

Beyond these more general obstacles, individual courts and justice agencies may also resist JIS integration. Courts and agencies may believe that their individual resources are too limited to participate in such a large project. Some might believe that the complexity of the justice process is so great that any attempt to automate and integrate it is doomed to end in failure. Others may be hesitant to rely on other courts’ and agencies’ technical staff for mission-critical computing and information. Still others may fear a reduction in their ability to serve their core constituencies. Finally, over the years, courts and agencies have invested a great deal of time, money, and other resources in developing electronic information systems that successfully address their core operational processes and requirements. Although these systems employ outmoded technologies, have limited functionality, cannot be adapted to emerging information processing needs, perpetuate islands of isolated data, and store information in data formats incompatible with other systems used by that court or agency, most courts and justice agencies are still comfortable with their systems. If these systems have limitations, they have developed work-a-rounds to compensate for them. The idea of having to reengineer their core business processes, or worse, to have another court or agency dictate to them how they must perform their core mission, is not a prospect most courts and justice agencies enthusiastically embrace.

To overcome these obstacles, the most promising strategy for developing an integrated justice information system is one that forswears any attempt to impose identical hardware and software solutions on the myriad courts and agencies to be integrated and focuses instead upon achieving interoperability and data exchange among the existing electronic information systems already in use. Interoperability and data

---

26. Such emerging information processing needs include the ability to handle new data formats, such as those required for storing digital mugshots, digital fingerprint records, and other images, to provide email notification to the parties of a case before the court, to include email as part of the case file, to permit the electronic filing of pleadings, motions, and other documents, to engage in the remote scheduling of court proceedings with parties, attorneys, law enforcement officers, and witnesses, and to provide public access to court documents over the Internet.
exchange, however, require standards—standards governing not only the communication of one computer system with another, but the translation of data from one format to another. Standards governing interoperability, data exchange, and data integrity would permit the creation of an open data-sharing network using middleware applications and data-warehousing solutions, while at the same time leaving in place the existing information systems of the courts, justice and public safety agencies, social services, and treatment providers that have been tailored to meet their unique operational requirements and the needs of their different constituencies.

Given the importance of standards to open data-sharing networks and, *a fortiori*, to the development of integrated justice information systems, it is hardly surprising that there are currently underway a number of efforts to develop various standards governing interoperability and data exchange tailored to the unique needs of the courts and the greater justice community. Two of these efforts, in particular, warrant our attention: the Global Justice XML Data Model being developed by the Office of Justice Programs in the United States Department of Justice (DOJ) and the LegalXML project of the Organization for the Advancement of Structured Information Standards (OASIS). The Global Justice XML Data Model is intended to facilitate “increased interoperability among and between justice and public safety information systems” and is “a significant milestone in the process of developing appropriate standards for expressing the baseline data needs of the justice and public safety communities and their related partners.” The LegalXML project produces, *inter alia*, standards for electronic court filing, court documents, legal citations, transcripts, and criminal justice intelligence systems. As the names of these two development

27. Despite “the significant efforts of government and industry to develop all manner of standards, additional standards governing technology, data integrity, and interoperability are still needed to help state and local agencies integrate.” NAT’L TASK FORCE ON COURT AUTOMATION & INTEGRATION, supra note 4, at 40. Recognizing this need, the National Task Force on Court Automation and Integration has advised that “communications protocols—such as TCP/IP (Transmission Control Protocol/Internet Protocol), frame relay, Internet/intranet standards, and universal transaction format standards like XML (Extensible Markup Language) that allow users to design the system the way they like—should be researched to determine their applicability to the development of integrated information systems.” Id. at 50.


efforts suggest, both the DOJ and OASIS have chosen XML as the language in which to develop their standards.

II. XML EXPLAINED

XML stands for Extensible Markup Language. Markup is information embedded in the text of a document that is not intended for printing or display. Experts generally agree that there are three kinds of markup: procedural, structural, and semantic. Procedural or presentational markup involves instructions for the display or printing of a document. It indicates “how” to render the text to which it is attached. The nonprintable codes embedded in a document to control the style or font of text by a word processing program are procedural markup—though such codes are more typically regarded as part of that word processing program’s proprietary file format. Structural or descriptive markup is used to indicate the type of text to which it is attached: a title, a quotation, headings, paragraphs, etc. It answers the question: what is this? Finally, semantic or content markup is used to indicate the meaning of a particular fragment of text. In a driver’s license, for example, semantic markup might indicate that a string of characters is the licensee’s name or that a string of numbers is the licensee’s birthday or identification number. In a trial transcript, semantic markup could be used to indicate when the speaker is the judge, a witness, or one of the attorneys, as well as whether the transcribed words are spoken in open court or at a sidebar. With the Global Justice XML Data Model and LegalXML projects, the DOJ and OASIS are both developing content markup for use by courts and other members of the justice and public safety communities.

Markup is associated with content and data using a markup language. A markup language is developed using a metamarkup language. The most powerful metamarkup language is Standard Generalized Markup Language (SGML). HyperText Markup Language (HTML), best known as the markup language used to create web pages on the World Wide Web, is a particular application of SGML—SGML was used to create it. Although SGML has been an international standard since 1986,
due to its complexity it cannot be used to develop more specialized markup language for use on the Web that addresses the needs of particular industries, professions, or other common interest communities. To develop specialized custom markup languages for use on the Web, a simpler metamarkup language was needed. Extensible Markup Language (XML), a subset of SGML, was developed to fill this need.\(^3\)

SGML and XML, as well as the particular markup languages created with them, all use tags to mark up content. Marked up or tagged content within a document is called an element. In particular markup languages such as HTML and LegalXML, each type of element is given a name. While elements can be nested, they cannot overlap, and all the elements composing a document must be contained within a single root or document element. Accordingly, an element’s content comprises either text or other elements. To make this a bit more concrete, consider the following simplified example\(^3\) of an element in a LegalXML document containing semantic markup used to tag data about a court filing stored in a court information system:

```
<courtFiling>
    <actor id="petitioner.1">
        <name>
            <personName>
                <namePrefix>Ms.</namePrefix>
                <firstName>Mary</firstName>
                <lastName>Smith</lastName>
            </personName>
        </name>
        <personDescription>
            <sex>female</sex>
            <birthDate>1972-06-15</birthDate>
            <maritalStatus>single</maritalStatus>
        </personDescription>
    </actor>
    <filingInformation id="filing.1">
        <courtInformation>
```

\(^3\) The new Web markup language known as XHTML is an application of XML. See infra note 46 and accompanying text.

\(^3\) A more complex example on which this example is based may be found at http://www.ncsconline.org/D_Tech/Standards/Documents/CourtDocument11-rev1/CourtDocument11/samples/petition_for_protection_filingCDATA.xml (last visited Jan. 9, 2004).
As this example shows, markup tags in markup languages created using XML begin with an opening angle bracket <, end with a closing angle bracket >, and include the name of the type of element of which they form a part. There are three kinds of markup tags in a markup language created using XML: start tags, end tags, and empty tags.  

Start tags are formed by placing angle brackets around the name of an element: <ElementName>. They indicate the beginning of an element of the kind named. End tags are formed from start tags by inserting a backward slash / between the opening angle bracket < and the element name for that tag: </ElementName>. An end tag indicates the end of the element whose beginning is indicated with the corresponding start tag. As the foregoing suggests, start and end tags form pairs, and in markup languages created using XML, one cannot place a start tag in a document without also using its corresponding end tag. A simple

example of marked up text conforming to these rules is:
\begin{verbatim}<ElementName>ElementContent</ElementName>.</verbatim>

In the simplified LegalXML example, the root element is called \texttt{courtFiling}. The root element envelops all the other elements of this example and extends from the start tag \texttt{<courtFiling>} to the end tag \texttt{</courtFiling>}. More specifically, the \texttt{courtFiling} element contains three elements: \texttt{actor}, \texttt{filingInformation}, and \texttt{leadDocument}. The \texttt{actor} element has nested within it the \texttt{name} and \texttt{personDescription} elements, which in turn contain structured information—another term for marked up content—about the filer’s first name, last name, sex, birth date, and marital status. The \texttt{filingInformation} element has nested within it the \texttt{courtInformation} and \texttt{caseInformation} elements, which in turn contain structured information about the physical location, type, and name of the court, as well as about the category, case number, year, and title of the case in which the document was filed. Note that although many elements are nested in others, no elements overlap: if an element’s start tag occurs between a pair of tags forming another element, then its end tag also occurs between that pair.

Unlike the \texttt{name} and \texttt{filingInformation} elements, the \texttt{leadDocument} element is an empty element. An empty element is used to place markup that contains neither text nor other elements. In our example, an empty element is used to represent the document filed because the actual document filed is not included in the structured information—the marked up content—actually presented. Empty elements may be represented using start tags that have a backward slash / inserted between the closing angle bracket > and the element name: \texttt{<ElementName/>}.\footnote{At the option of the markup’s author, a space may be inserted between the element name and the backward slash: \texttt{<ElementName/>}.} In our example, the actual document filed is associated with the empty tag \texttt{<leadDocument/>} by assigning it a unique document ID and placing that ID number in the tag as an attribute.

Attributes permit additional information concerning elements to be included in tags. Only start and empty tags may include attributes. Among other things, attributes enable the markup author to give each element in a document a unique identifier, to associate URLs with text or images in order to create hyperlinks, as well as to insert applicationspecific instructions for processing the structured information and data.
In our example, the *leadDocument* element included the attribute: id="doc001.01-2-13059-5-SEA". As you can see, attributes are formed by concatenating an attribute’s name to an equal sign followed by the attribute’s value—the information that the attribute is intended to convey—within quotation marks. They are inserted into markup tags between the element name and the closing angle bracket: <ElementName AttributeName="AttributeValue">. In the simplified LegalXML example above, the *actor*, *postalAddress*, and *caseInformation* elements, among others, also contained attributes.

The first and current version of XML developed by the XML Working Group is referred to as XML Specification 1.0 (XML 1.0). This specification comprises the syntax for creating well-formed XML documents as well as the rules for creating new markup tags. Using XML 1.0 as a metamarkup language, one can create markup languages such as LegalXML that not only facilitate data exchange by specifying the meaning or significance of industry-specific data, but also specify “the structure of that data and how various elements are integrated into other elements.” The definition of new markup tags as well as the rules and constraints governing their use are expressed in machine-readable documents called Document Type Definitions (DTDs) and XML schemas. The DOJ’s Global Justice XML Data Model project, for example, employs XML schemas to define the approximately 2000 unique data elements included in its XML standard for use by the justice

---

37. Attributes are also used in HTML. For example, to create a hyperlink in a document using HTML, one must not only use a pair of anchor tags to surround the specific text, which when clicked will trigger the web browser to jump to another location on the Internet, but one must also include the URL of the target location. This additional information is conveyed using the Hypertext REFerence attribute, HREF: <A HREF="http://www.targetLocation.com">Hyperlink Text</A>.  
38. S. LAURENT, supra note 35, at 46.  
39. Id.  
41. Id.  
43. While XML schemas can be used instead of DTDs to govern the use of customized markup tags, they are the subject of a separate W3C specification. See World Wide Web Consortium, *XML Schema Part 1: Structures*, at http://www.w3.org/TR/xmlschema-1/ (last visited Sept. 11, 2003). The central difference between a DTD and XML schema is that an XML schema is itself an XML document, while a DTD is not. Moreover, XML schemas are more powerful and precise than DTDs. For example, XML schemas allow for data typing, but DTDs do not. For these reasons, XML developers generally prefer using XML schemas to DTDs when defining new markup tags.
and public safety communities. OASIS’ LegalXML project, by contrast, used a DTD for its Electronic Court Filing 1.1 Proposed Standard.

An XML standard, such as those being developed in the Global Justice XML Data Model and LegalXML projects, is simply an application of XML that includes various markup tags, Document Type Definitions, or XML schemas tailored to the needs of a particular industry or user community. XML 1.0 is used to develop XML standards. For example, the World Wide Web Consortium (W3C) used XML 1.0 to develop XHTML 1.0, an XML application intended to replace HTML for use on the Web. XHTML 1.0 contains all of the markup tags of HTML as well as a DTD defining the structure of a valid Web page, i.e., HTML document. XML 1.0 has also been used to develop standards for use in banking, finance, telecommunications, education, and myriad kinds of businesses. Documents that comply with XML standards are well-suited for use over the Internet, easily understood by humans, and machine-readable. It is hardly surprising, therefore, that the DOJ and OASIS both chose XML to develop standards for use by the courts, justice and public safety agencies, social services, and treatment providers that form the justice and public safety communities.

As the framework and metalanguage within which standards are developed, XML “is the ‘glue’ that promotes interoperability—it allows systems already in use and those being developed to communicate with each other and paves the way for future expanded collaboration between agencies.” Data structured using XML markup is not tethered to particular vendors, original equipment manufacturers (OEMs), operating systems, applications, communication protocols, or storage media and “can be shared between different systems, up and down the levels of...”

47. HOLZNER, supra note 42, at 10.
Rise of the Machines

agencies, across the nation, and around the world, with the ease of using the Internet."49

XML promotes interoperability by enabling the creation of XML standards tailored to the needs of different industries and user communities. In particular, XML standards tailored to the operational requirements of the courts and other members of the justice and public safety communities promote interoperability by defining elements that can express the types of data used by courts and other members of the justice and public safety communities as structured information. Using such elements, not only can data from any justice information system be re-expressed as structured information, but structured information conforming to the standard can be translated into the data format of any justice information system. As a result, XML can be used to create applications and standards that bridge the information chasm separating any justice information systems using incompatible data formats. One need only translate the data from one justice information system into a document conforming to an appropriate XML standard and then translate that document into the data format of the other justice information system. Since any type of data or information covered by an XML standard can be translated from or to structured information conforming to that standard, such structured information forms a bridge between any two justice information systems.50 In this fashion, data and messages can be moved from one justice information system to any other with XML markup acting as a lingua franca or intermediate link between different and incompatible justice information systems.

By using XML to develop XML standards tailored to the operational requirements of the courts and other members of the justice and public safety communities, interoperability and data exchange among disparate and incompatible information systems can be achieved and the obstacles to an integrated justice information system surmounted. In light of the


50. This process is generally automated using software running on application servers connected to the communicating justice information systems. The justice information system sending data passes the data in its native data format to an application server that translates that data into XML elements defined by the appropriate XML standard. This application server then transmits the resulting XML elements to another application server that translates this structured information into the data format of the target justice information system. The reformatted data is then passed to the target information system. When required, the target information system can respond to the sending justice information system by reversing this process.
extensive benefits arising from the integration of justice information systems, it is inevitable that the courts and other members of the justice and public safety communities will migrate to case management information systems that can output XML structured data and information either directly or through a middle tier application server.

III. DOCUMENTS DISSOLVED: FROM MANILA FOLDERS TO STRUCTURED INFORMATION

At least since the construction of permanent courthouses, individuals—attorneys in particular—have generally conveyed information to the courts in the form of paper documents: petitions, bills, complaints, answers, motions, memoranda, affidavits, attachments, transcripts, requests for findings of fact and conclusions of law, informations, indictments, notices of appeal, and briefs. Courts too have typically responded with paper documents: written opinions, judgments, decrees, findings, minutes, injunctions, writs of mandamus and prohibition, as well as other kinds of orders. Indeed, some judges have created documents within documents by inscribing their rulings—granted or denied—as marginalia on motions submitted by the parties. In most courthouses, these paper documents are sorted by case, collected together in manila folders and expanding Pendaflex® or Oxford® file pockets, and stored in a filing room that provides ready access to court clerks and their assistants.

While at one time unavoidable, operating a court system in which paper documents mediate the exchange of information among judges, court personnel, litigants, attorneys, witnesses, the press, and the interested public is highly inefficient. Among the many costs incurred by parties, their attorneys, court personnel, and the quality of justice generally, we may highlight just a few. To begin with, filing a paper document with a court requires that one travel to the courthouse during court business hours, hand the document to a clerk, and receive a paper time-stamped receipt. Not only does this process increase the cost of access to the justice system, but when performed by an attorney, it prevents a more productive use of his or her time. Once filed with the court, court personnel must update the docket and physically file the document in the filing room. This not only lessens the productivity of court personnel but also creates the opportunity for data entry and filing errors. When needed again, paper documents are cumbersome, difficult to retrieve quickly, and require the user’s physical presence at the courthouse. Moreover, only one person can read a paper document at a
Rise of the Machines
time. This limitation leads to restricted access when judges and judicial clerks are working with case files as well as costly copying. Additionally, when paper documents and case files are moved among different users, they are not infrequently lost, disassembled, or misfiled—reducing access still more. Paper documents are also notoriously difficult to keep up-to-date. As a result, judges, court personnel, attorneys, the press, or the public may find themselves inadvertently relying on a document containing outdated information. In light of these difficulties and disadvantages, it is hardly surprising that as information and communication technologies have evolved, virtually everyone connected with the justice system has embraced the search for an electronic alternative to paper documents.

Within the federal courts, the search for an electronic alternative to paper documents may be dated from the 1991 amendment of Rule 25 of the Federal Rules of Appellate Procedure to permit electronic filing. As amended, Rule 25 provides that

[a] court of appeals may by local rule permit papers to be filed, signed, or verified by electronic means that are consistent with technical standards, if any, that the Judicial Conference of the United States establishes. A paper filed by electronic means in compliance with a local rule constitutes a written paper for the purpose of applying these rules.51

Five years later, similar changes were made to the Federal Rules of Civil Procedure52 and the Federal Rules of Bankruptcy Procedure.53 The Federal Rules of Criminal Procedure state simply that “[a] paper must be filed in a manner provided for in a civil action”—thereby allowing electronic filing by incorporation.54 Since 2001, the Federal Rules of Civil Procedure also permit service by electronic means.55 The Federal Rules of Appellate Procedure followed suit in 2002.56 Significantly, both also permit service by electronic means using the court’s own transmission equipment.57 Thus, in only twelve years, the federal judiciary has gone from permitting U.S. courts of appeals to accept

52. FED. R. CIV. P. 5(e).
53. FED. R. BANKR. P. 5005(a)(2).
54. FED. R. CRIM. P. 49(d).
55. FED. R. CIV. P. 5(b)(2)(D), 5(b)(3).
56. FED. R. APP. P. 25(c)(1)(D).
57. See FED. R. CIV. P. 5(b)(2)(D); FED. R. APP. P. 25(c)(2).
electronic filings to allowing U.S. district and circuit courts not only to accept service made by electronic means but to facilitate such service using its own equipment.

Just as the rules authorizing electronic filing have evolved from the early 1990s to the present, so has the technology. In its earliest implementation, electronic filing meant “filing by fax.” A paper document was inserted into a fax machine at one end, an electronic signal was transmitted over the telephone lines to the courthouse fax machine, and a paper document came out at the other end. The court clerk then filed the faxed document as he or she would any other paper document. In 1995, commercial traffic was permitted onto the Internet and soon thereafter “filing by fax” was replaced by “filing by email attachment.” Lawyers could now simply attach their WordPerfect, Microsoft Word, or Adobe Portable Document Format (PDF) documents to an email addressed to the court clerk and with the click of a button the filing would be at the courthouse. At first, court personnel printed the documents received via email and then processed them as they would any paper document. As courts reengineer their work flow to benefit from the lower costs of electronic storage, however, many courts are now storing all their documents in electronic form—scanning the paper documents that they receive and converting them to TIFF image or Adobe PDF files.

The benefits and savings that result from electronic filing have been recognized for a long time. In 1997, the Administrative Office of the United States Courts created a comprehensive catalog of benefits provided by electronic filing that included the following: immediate access to court documents in the courtroom; simultaneous access to the same document by many individuals; twenty-four-hour access to court documents inside and outside the courthouse; reduced file-handling and maintenance; reduced staff time for tasks such as pulling and reshelving files and making copies for the public; simplified archiving and file retrieval; easy and quick transfer of case files among courts, chambers, and other court units; elimination of misfiled papers; reduced time for dictation and retyping as portions of one document can be easily transferred to another using the cut-and-paste operation of word processing software; reduced need for physical space in the courthouse to store documents; elimination of the need for archival storage using microfilm; reduced data entry errors through automated docketing of electronically filed documents; reduced need to assist with public access to documents; reduced need for filing duplicate paper copies of
documents by attorneys as well as for the handling of such copies by court personnel; greater file integrity and security using validity checks; remote viewing of court documents over the courthouse network or Internet; enhanced ability to share, annotate, and edit documents through email; and the ability to perform full text searches within individual documents and across an entire file system.\textsuperscript{58} In light of the substantial advantages and savings provided by electronic filing, it is hardly surprising that both federal and state courts are embracing electronic filing through pilot programs and permanent implementations.

From a systems integration perspective, electronic filing also permits a court’s document management system to be integrated with its case management system. Integrating these two systems, for example, allows a user viewing an electronic docket to access and read a particular document by clicking on that document’s title.\textsuperscript{59} Moreover, just as XML can be used to integrate case management information systems across courts and other justice and public safety agencies, XML can also be used to facilitate interoperability and data exchange among integrated case management and electronic case file systems. Using an appropriate XML standard, an electronic document is merely another form of content to be marked up using tags. For example, using a pair of \texttt{documentContent} tags, one could simply insert an Adobe PDF or Microsoft Word document as an element in a LegalXML message being transmitted to another information system.

More significantly still, one may use tags defined using an appropriate XML standard to mark up different kinds of information in electronic documents filed with the courts. Tagging key pieces of information in court documents would permit court information systems to read, understand, and manipulate this information without human intervention. For example, if the plaintiffs’ names, the names of their attorneys, and their attorneys’ phone numbers were all tagged using LegalXML in electronic court documents containing this information, then software could be developed that could create—automatically or on


request—a list of the plaintiffs’ attorneys and their phone numbers for any case. Even more importantly, software could be developed that creates copies of court documents with certain tagged categories of information redacted. Sensitive personally identifying information, for example, could be removed from court documents when accessed by the public. Recognizing the need for authors to create electronic documents that include XML markup, upcoming releases of both Microsoft Word and Adobe Acrobat—the application used to create Adobe PDF documents—will include support for XML.60

With the development of XML standards, documents are becoming anachronistic. By tagging all the information contained in a court document, it is possible to dispense with documents altogether—through dissolving them into structured information. After all, a document is only a particular view of the information that it contains. Rather than being restricted to one particular view of that data, using structured information one could select or create a view of the data optimized for the task to which that data is relevant. Imagine, for example, being able to display simultaneously the conflicting factual claims contained in a plaintiff’s complaint and a defendant’s answer, or an argument and its critique culled from one side’s memorandum in support of a motion and the other side’s memorandum in opposition. Such tailored views of case data as well as traditional documentary views could easily be created if we filed structured information with the courts rather than documents—electronic or paper.61 Admittedly, however, judicial information systems for filing structured information are still somewhat in the future.

IV. THE QUESTION OF PUBLIC ACCESS TO COURT RECORDS OVER THE INTERNET

Already in some jurisdictions—and soon in most—courts are integrating their case management information system with an electronic docket and document management system. Using XML standards tailored to the operational requirements of the courts and other members


61. Different views of XML tagged data are created using Cascading Style Sheets (CSS) or Extensible Style Language (XSL). Using Extensible Style Language: Transformations (XSLT), one can even create and save documents in particular formats—such as Adobe PDF or Microsoft Word—out of structured information.
Rise of the Machines

of the justice and public safety communities, these integrated information systems are able to interoperate and exchange data with the information systems of other courts, justice and law enforcement agencies, social services, and treatment providers. Incompatible data formats, communications protocols, and application programming interfaces are obstacles surmounted. Truly integrated justice information systems at the state, regional, and national levels are within reach. Also within reach is public access to court records on a scale and with an ease never before imagined.

Just as XML permits the myriad, incompatible information systems of the courts and other members of the justice and public safety communities to interoperate and exchange data, this same technology enables court information systems to interoperate and exchange data with personal computers connected to the Internet. As a result, it is now technically possible to permit members of the public to access court data and case files from the privacy and comfort of their own home or office, in real time, twenty-four hours a day, seven days a week. Moreover, to the extent that courts are already modifying existing information systems or designing new ones to output structured information and electronic documents as XML messages, this access can be granted to the public without any additional development costs.

Recognizing that it is now technically possible to provide public access to court records over the Internet leads inexorably to the question whether the courts ought to provide such access and, if so, to what extent. At least three answers to this question are possible. One might believe that providing public access to case files over the Internet would do too great a harm to the privacy of litigants and others discussed in court documents and, therefore, conclude that the public ought not to be permitted access to court documents over the Internet. By contrast, one might observe how public access to the courts has been an enduring and fundamental value in our society[62] and conclude that the public should be able to access court files and documents over the Internet to the same extent that they can access them in the courthouse—that courts should not discriminate with respect to methods of access. Finally, one might

balance the importance of public access to the courts against the privacy interests of individual litigants and conclude that while the public should have some access to court files and documents over the Internet, it should be more limited than the access to case files provided in the courthouse.

Of these three positions, I embrace the second: I do not believe that courts should discriminate with respect to methods of access. For the reasons explained in Parts V and VI, especially as courts migrate to XML-based justice information systems, I do not believe that limiting public access to court documents over the Internet is necessary to protect the legitimate privacy interests of litigants or others before the court. To the extent that information should be withheld from the public because of legitimate concerns about public safety or privacy, such information should be inaccessible to the public at the courthouse as well.

An alternative position has been taken by the Conference of Chief Justices (CCJ) and the Conference of State Court Administrators (COSCA). Endorsing the position that public access to court records over the Internet ought to be more limited than public access to these records in the courthouse, the CCJ and COSCA have published a set of model guidelines for granting public access to court records (CCJ/COSCA Guidelines or Guidelines). The CCJ/COSCA Guidelines include definitions of both “public access” and “court record.” Under the Guidelines, public access “means that the public may inspect and obtain a copy of the information in a court record,” while a court record is defined to include “[a]ny document, information, or other thing that is collected, received, or maintained by a court or clerk of court in connection with a judicial proceeding” as well as “[a]ny index, calendar, docket, register of actions, official record of the proceedings, order, decree, judgment, minute, and any information in a case management system created by or prepared by the court or clerk of court that is related to a judicial proceeding.”

64. Id. at 12 (CCJ/COSCA Guideline § 3.10—Definition of Court Record), 17 (CCJ/COSCA Guideline § 3.20—Definition of Public Access).
65. Id. at 17 (CCJ/COSCA Guideline § 3.20—Definition of Public Access).
66. Id. at 12 (CCJ/COSCA Guideline § 3.10—Definition of Court Record). Currently, disagreement exists within the justice community over whether administrative records of a court should be included within the definition of a court record. The Guidelines have sidestepped this.
Under the Guidelines’ general access rule, information in court records should be accessible to the public unless its release is prohibited by federal or state law or the court record is sealed by the court.\textsuperscript{67} This general access rule applies to all court records—both paper and electronic.\textsuperscript{68} The degree of access recommended by the Guidelines, however, varies with the method of access used. The Guidelines distinguish courthouse access from remote access. Remote access is defined as “the ability to electronically search, inspect, or copy information in a court record without the need to physically visit the court facility where the record is maintained” and, therefore, would include access to court records over the Internet.\textsuperscript{69} Significantly, the Guidelines do not recommend that the degree of access provided remotely be as extensive as the degree of access provided at the courthouse. In particular, while public access to case files and documents must be provided at the courthouse, remote access to case files and documents is not recommended.\textsuperscript{70} Remote access is presumed only for: litigant indexes filed with the court; listings of new case filings, including the names of the parties; registers of actions showing what documents have been filed in a case; calendars or dockets of court proceedings; judgments, orders, or decrees; and liens affecting real property.\textsuperscript{71} Explaining its decision to recommend limiting the public’s remote access to this type of information, the commentary to the Guidelines notes that “[t]he summary or general nature of the information is such that there is little risk of harm to an individual or unwarranted invasion of privacy or proprietary business interests.”\textsuperscript{72}

In contrast to the CCJ and the COSCA, and more closely aligned to my own view, the Judicial Conference of the United States Committee on Court Administration and Case Management on Privacy and Public Access to Electronic Case Files (Committee) has taken the position with regard to civil case files that with one exception, the courts should not discriminate among modes of access: “the federal courts recognize that controversy by allowing but not requiring the definition of court record to include administrative records and information. See id. (CCJ/COSCA Guideline § 3.10(a)(3)).

\textsuperscript{67} Id. at 23, 45, 53 (CCJ/COSCA Guidelines §§ 4.00, 4.60, 4.70(a)).
\textsuperscript{68} Id. at 22 (CCJ/COSCA Guideline § 4.00—Applicability of Rule).
\textsuperscript{69} Id. at 19 (CCJ/COSCA Guideline § 3.30—Definition of Remote Access).
\textsuperscript{70} See id. at 27 (CCJ/COSCA Guideline § 4.20—Court Records in Electronic Form Presumptively Subject to Remote Access by the Public).
\textsuperscript{71} Id.
\textsuperscript{72} Id.
the public should share in the benefits of information technology, including more efficient access to court case files.”

Accordingly, while the Committee recognizes that “[a]s a practical matter, during this time of transition when courts are implementing new practices, there may be disparity in access among courts because of varying technology,” it emphasizes that with one exception, the degree of access provided over the Internet should be the same as the degree of access provided at the courthouse:

documents in civil case files should be made available electronically to the same extent that they are available at the courthouse with one exception (Social Security cases should be excluded from electronic access) and one change in policy (the requirement that certain ‘personal data identifiers’ be modified or partially redacted by the litigants).

Similar recommendations for electronic access are made with respect to bankruptcy and appellate case files. Only with regard to criminal case files does the Committee recommend restricting public access over the Internet. Even in civil cases, however, the Committee recognizes that “[c]ertain types of cases, categories of information, and specific documents may require special protection from unlimited public access.” Thus, its recommendation contemplates that certain personal

73. Judicial Conference Comm. on Court Admin. & Case Mgmt., Report of the Judicial Conference Committee on Court Administration and Case Management on Privacy and Public Access to Electronic Case Files (2001), available at http://www.privacy.uscourts.gov/Policy.htm. It should be noted that the Committee defines case files more narrowly than the CCJ/COSCA Guidelines. Under the Committee’s definition, [t]he term “case file” (whether electronic or paper) means the collection of documents officially filed by the litigants or the court in the context of litigation, the docket entries that catalog such filings, and transcripts of judicial proceedings. The case file generally does not include several other types of information, including non-filed discovery material, trial exhibits that have not been admitted into evidence, drafts or notes by judges or court staff, and various documents that are sometimes known as “left-side” file material.

Id.

74. Id.

75. Id.

76. Id. The Committee recommended that “public remote electronic access to documents in criminal cases should not be available at this time [September 2001], with the understanding that the policy will be reexamined within two years of adoption by the Judicial Conference.” Id. It subsequently amended this recommendation on criminal cases to allow two exceptions: one for high profile criminal cases, and the other for pilot programs granting online access to documents in criminal case files. See Judicial Conference Comm. on Court Admin. & Case Mgmt., Limited Exceptions to Judicial Conference Privacy Policy for Criminal Case Files (2002), available at http://www.privacy.uscourts.gov/amend.htm.

77. Judicial Conference Comm. on Court Admin. & Case Mgmt., supra note 73.
data identifiers—Social Security numbers, dates of birth, financial account numbers, and names of minor children—“will not be included in its full and complete form in case documents, whether electronic or hard copy.”

In the E-Government Act of 2002, the United States Congress also endorsed public access to court documents over the Internet. Section 205(a) of the E-Government Act requires that the United States Supreme Court as well as all federal circuit, district, and bankruptcy courts establish and maintain a website that provides, *inter alia*, “access to documents filed with the courthouse in electronic form, to the extent provided under subsection (c).” Subsection (c) requires all electronic documents that are publicly available at a federal courthouse to be made publicly available online as well. It also requires the U.S. Supreme Court to promulgate uniform rules to protect privacy and security concerns relating to electronically filed documents and their public availability under this section. “To the extent that such rules provide for the redaction of certain categories of information,” subsection (c) permits a party to file an unredacted copy of the document under seal, which a court may, in its discretion, accept in lieu of or in addition to a redacted copy in the public file. Although the federal courts might question Congress’ constitutional authority to prescribe the manner in which the federal courts must provide public access to case files and documents, it is clear that to the extent feasible and with few exceptions, both the federal courts and Congress are committed to making electronic documents publicly available online to the same extent that they are publicly available at the courthouse.

I wholly endorse this policy’s general commitment to the even-handed treatment of paper and electronic documents. Unlike the Committee, however, as explained in Parts V and VI, I believe that the same even-handed treatment of case files and documents can be

---

78. *Id.*


80. *Id.* § 205(a). This section also requires that the website provide access to docket information and, to the extent feasible, “to post online docket sheets with links allowing all filings, decisions, and rulings in each case to be obtained from the docket sheet of that case.” *Id.* § 205(d).

81. *Id.* § 205(c)(1). Electronic documents include electronically filed documents as well as paper documents filed with the court of which the court has created an electronic version by scanning or other means.

82. *Id.* § 205(c)(3)(A)(i).

83. *Id.* § 205(c)(3)(A)(iv).
extended to criminal and Social Security cases while respecting the legitimate public safety and privacy interests of litigants and others involved in those cases.

At the present time, the apparent trend among state courts is to follow the recommendations of the CCJ/COSCA Guidelines rather than the policies of the United States Congress and the Committee. For example, in Massachusetts, the Supreme Judicial Court has “concluded, at least initially, that an intermediate level of access to court information is appropriate on the Web, one that provides less information than is available at a courthouse.”\(^8^4\) Accordingly, its policy governing public access to court records over the Internet “does not allow documents submitted to a court in connection with a case to be published on a Court Web site.”\(^8^5\) Moreover, while its policy does permit public access over the Internet “to docket and calendar information that is or will be maintained in computerized case management systems,” the Supreme Judicial Court has emphasized that “the law does not require courts to provide electronic access to court case information” and expressly prohibits publishing any criminal case information on the Web that “would identify a specific criminal defendant by name.”\(^8^6\) The Supreme Judicial Court has explained that it adopted this policy of limited public access over the Internet to address “concerns about the substantial intrusions into privacy interests that could accompany publication of personal information on a Court Web site.”\(^8^7\)

Another state that appears to be following the CCJ/COSCA Guidelines is California. Rule 2073 of the 2003 California Rules of Court governing public access to trial court records requires that “[a] court that maintains . . . records in electronic form must provide electronic access to them at the courthouse, to the extent it is feasible to do so, but may provide remote electronic access only to the records governed by (b)(1),” where the records governed by subsection (b)(1) include only the “[r]egister of actions (as defined in Gov. Code, § 69845), calendars, and indexes.”\(^8^8\) Remote access to all other case records, including all documents filed in the case, is prohibited. Again

\(^8^4\) Supreme Judicial Court, Policy Statement by the Justices of the Supreme Judicial Court Concerning Publication of Court Case Information on the Web 1 (May 2003), at http://www.state.ma.us/courts/web-pubpolicy.pdf.
\(^8^5\) Id.
\(^8^6\) Id. at 1–2.
\(^8^7\) Id. at 1.
\(^8^8\) CAL. CT. R. 2073(b)(1), (c) (2003).
Rise of the Machines

the principal motivation for restricting the court records to which the public may have remote access appears to be a concern over privacy: “[t]he rules in this chapter are intended to provide the public with reasonable access to trial court records that are maintained in electronic form, while protecting privacy interests.” 89

In contrast to many state courts, the federal courts are phasing in a new case management/electronic case files (CM/ECF) system that implements the recommendations of the Committee. This system not only permits “litigants to file and access documents electronically via the Internet, 24 hours a day and 7 days a week,” but also provides public access to these electronic case files by the same means. 90 The United States District Court for the Southern District of New York, for example, has developed an extensive website to support its own implementation of this system that includes a publicly accessible online tutorial for attorneys and law firm staff. 91 These systems are already being used in twenty-five district courts, sixty bankruptcy courts, the Court of International Trade, and the Court of Federal Claims. “Under current plans, the number of CM/ECF courts will increase steadily each month into 2005” with implementation for appellate courts scheduled to begin in late 2004. 92

The foregoing suggests that the CCJ/COSCA Guidelines and the state court policies and rules that embody them restrict public access to the court records over the Internet due to a general sense of dread and anxiety over possible harms to the privacy interests of litigants and others who come before the courts. This anxiety over privacy, however, is completely misplaced. On the one hand, courthouse access provides no greater protection to privacy than access over the Internet. Any journalist can access information in case files at the courthouse and publish it on the front page of his or her newspaper. That most litigation is dull as dishwater is no response. Credit bureaus and other information aggregators will gladly mine the most boring case files for nuggets of

89. Id. 2070(a).
91. The court’s ECF website is available at http://www.nysd.uscourts.gov/cmecf/cmecfindex.htm (last visited Jan. 9, 2004), while the online tutorial may be found at http://www.nysd.uscourts.gov/cmecf/training/ecf100/index.html (last visited Jan. 9, 2004).
data and pass the cost on to their information consumers. Indeed, one of the reasons that the Committee decided against limiting public access to court records over the Internet was because it would simply foster a cottage industry of court data resellers—making the resellers rich, while depriving the public of access and leaving privacy unprotected. On the other hand, as explained in Part V, providing access to court records and documents over the Internet, if done properly using XML-based justice information systems, can actually protect sensitive information better than restricting public access to paper case files in a courthouse.

At the present time, the CCJ/COSCA Guidelines—and the court policies and rules that implement them—deny the public access to court documents over the Internet because some documents may include sensitive information, while those same documents are made available in full for anyone to read at the courthouse. As a result, the public is denied easily accessible information that it should have, while it is given sensitive information that it should not. The problem is that the present information systems cannot distinguish between private and public forms of information within a document. However, as justice information systems are developed around structured information rather than generic documents and files, court information systems will be able to intelligently process different forms of information, preventing potentially harmful disclosures while simultaneously permitting the same degree of public access at the courthouse and over the Internet.

V. SEALING INFORMATION TO PROMOTE PUBLIC SAFETY AND PROTECT PERSONAL SECURITY

The most compelling reason for restricting public access to court records is not privacy but public safety—protecting the physical, psychological, and economic security of individuals. The information in court records and documents is not all benign, and our commitment to open judicial proceedings and public access to court records has involved certain costs arising from the misuse of this information. Public access to information in court records can facilitate blackmail, extortion, stalking, sexual assault, subornation of perjury, identity theft, and fraud—to name only a few of the crimes that can be facilitated through the misuse of information obtained from court records. Fortunately, as court information systems come online that can accept, process, and

93. JUDICIAL CONFERENCE COMM. ON COURT ADMIN. & CASE MGMT., supra note 73.
output XML tagged structured information, information likely to facilitate such crimes can be withheld from the general public, and these social costs obviated without in any way diminishing the benefits of unrestricted public access to court records and documents at the courthouse and over the Internet.

The kind of information most subject to misuse is personal information: information or data linked to or identified with a particular person. While any fact can be associated with a particular individual, certain categories of personal information render a person particularly vulnerable to malfeasance and harm: these include a person’s address, telephone number, Social Security number, driver’s license identification number, bank accounts, debit and credit card numbers, and personal identification numbers (PINs). One of the most publicized cases of harm facilitated by information culled from public records involved the actress Rebecca Schaeffer, the star of My Sister Sam, a popular television series during the 1980s. As Representative James P. Moran recounted on the floor of the House of Representatives, “[a]lthough she had an unlisted home number and address, Ms. Schaeffer was shot to death by an obsessed fan who obtained her name and address through the [California Department of Motor Vehicles (DMV)].” Representative Moran also described how “[i]n Iowa, a gang of thieves copied down the license plate numbers of expensive cars they saw, found out the names and addresses of the owners [from the DMV] and robbed their homes at night.” On the floor of the Senate, Senator Barbara Boxer described an equally disturbing account of a thirty-one-year-old man in California who “copied down the license plate numbers of five women in their early twenties, obtained their home address from the DMV and then sent them threatening letters at home.” Senator Boxer read two of the letters into the Congressional Record. One read: “I’m lonely and so I thought of you. I’ll give you one week to respond or I will come looking for you.” The other read: “I looked for you though all I knew about you was your license plate. Now I know more and yet nothing. I know you’re a Libra, but I don’t know what it’s like to smell your hair while I’m kissing your neck and holding you in my arms.”

95. Id.
97. Id.
98. Id.
The information that facilitated these wrongful acts came from public records at state departments of motor vehicles and led to the passage of the Driver’s Privacy Protection Act of 1994, but the type of information used—an individual’s name and address—might just as well have come from court records accessed at a local courthouse.

Identity theft is another example of wrongdoing that can be facilitated by information culled from court records. “An identity thief co-opts some piece of your personal information and appropriates it without your knowledge to commit fraud or theft.” An identity thief may call a person’s credit card issuer, change the billing address by impersonating that person and then run up charges on that account, or the identity thief may simply open a new credit card account in that person’s name. In either case, when the charges remain unpaid, the delinquent account is reported on the victim’s credit report. An identity thief may also open bank accounts in a victim’s name and proceed to write bad checks. Telephone and wireless accounts as well have been established in victims’ names. Identity thieves have also been known to purchase automobiles by taking out auto loans in the name of a victim. Some identity thieves have even filed for bankruptcy in the name of a victim to avoid eviction or to discharge debts that they incurred in a victim’s name. Perhaps the most egregious form of identity theft occurs when an identity thief is arrested and gives a victim’s name as his or her own: when the thief fails to appear on a scheduled court date, the arrest warrant is issued in the name of the victim. The Federal Trade Commission (FTC) estimates that in the United States over the last year alone losses from identity theft have approached fifty billion dollars. The FTC also estimates that over the last year approximately ten million people in the United States discovered that they were victims of identity theft and that each victim spent on average five hundred dollars and thirty hours resolving the resulting problems.

Personal information that facilitates these kinds of wrongs should not be accessible to the public either at the courthouse or over the Internet. It places an individual in jeopardy of physical, psychological, and economic harm without furthering any of the benefits of public access to

102. Id. at 7.
court records. The principal benefits promoted by public access to court records may be grouped into two categories: accountability and citizen education. With respect to the former, public access to court records makes the courts themselves directly accountable to the people. As Justice Oliver Wendell Holmes noted over one hundred years ago:

[i]t is desirable that the trial of causes should take place under the public eye, not because the controversies of one citizen with another are of public concern, but because it is of the highest moment that those who administer justice should always act under the sense of public responsibility, and that every citizen should be able to satisfy himself with his own eyes as to the mode in which a public duty is performed.\(^\text{103}\)

Reviewing court records not only achieves this end and promotes public trust and confidence in the judiciary, but it also demonstrates to the public that the rule of law is being upheld and the law enforced. Moreover, insofar as government agencies, corporations, businesses, charities, politicians, and other individuals in whom a public trust has been invested come before our courts, public access to court records also keeps them accountable to the people. With respect to citizen education, public access to court records teaches the general public how the courts operate and informs them of the results of cases as well as the evidence that supports those results. By educating the public in this fashion, public access to court records promotes a set of stable and predictable rules by which we can govern ourselves and instructs members of the public with respect to people, circumstances, and business practices that might cause them harm.

Withholding sensitive personal information from the public when they access court records will neither undermine nor subvert any of these benefits. The adjudicatory facts upon which a court relies to dispose of a case or controversy according to the rule of law need never include the specific, arbitrarily assigned street address of a person’s home, the precise series of numerals composing his or her telephone number, or the exact digits of his or her Social Security number. That a person has a Social Security number may be relevant to the just and rational disposition of a case, but the specific number will not be. That a person resides along a particular street or next to one of the parties may be relevant, but the exact house number will not be. Similarly, the general education that an individual might be expected to acquire from the

---

perusal of court records does not include committing to memory the street addresses of fellow citizens, their Social Security numbers, or their bank accounts. Accordingly, such information should be omitted from publicly accessible court records and documents, irrespective of their form or the public’s method of accessing them.

One must be careful, however, not to throw the baby out with the bath water. While personal information that renders an individual vulnerable to physical, psychological, and economic harm should not be released to the public, the information that surrounds personal information should be publicly accessible. Too often, information that would further accountability and citizen education is kept from the public because it cannot be effectively disentangled from personal information in documents or files that courts order sealed. While redacting the personal information contained in these documents and files is possible, it is quite costly, and many courts lack sufficient resources to do so on a regular basis.104 As a result, the public is denied access to information that it should have. Thus, courts are caught between the horns of a dilemma: either they deny the public access to information it should have or grant it access to information that it should not.

To pass between the horns of this dilemma, some jurisdictions have adopted court rules that permit parties to omit sensitive personal information from publicly accessible documents filed with the court. When exercising this option, parties submit their personal information on confidential court forms that are not publicly accessible. General Rule 22(g) of the Washington State Court Rules, governing access to family court records, provides a good example. Subsection 22(g)(1) states, in relevant part, that “[i]nformation filed by a party in any file or record . . . shall be available to the public unless . . . access is restricted under section (c)(2).” 105 Subsection (c)(2) provides a list of forms for collecting personal information—such as a party’s residence address, Social Security number, driver’s license number, telephone number, Social Security number of a child or date of birth of a child—that the public is restricted from accessing.106 An official comment to the rule

104. STEKETEE & CARLSON, supra note 63, at 18.
105. WASH. CT. GEN. R. 22(g)(1).
106. Id. 22(c)(2). Pursuant to General Rule 22(g)(3), “[a]ny person may file a motion, supported by an affidavit showing good cause, for access to any document otherwise restricted under section (c)(2).” Id. 22(g)(3). Moreover, pursuant to General Rule 22(g)(2), “[t]he parties may stipulate in writing to allow access to the public to any files or records otherwise restricted under section (c)(2).” Id. 22(g)(2).
Rise of the Machines

notes that if a party files a document containing sensitive personal information normally collected on the confidential forms listed in subsection (c)(2), “such documents shall be publicly available in the case record.”

Although paper and electronic forms are reasonable short-term measures for resolving this dilemma, once courts implement information systems that can accept, process, and output XML markup, a more efficient and versatile approach would use XML tags to mark up sensitive personal information and control access to it programmatically. For example, when a party or attorney authors a document for filing with a court, he or she could include markup for recognized categories of personal information that the public should be restricted from viewing. Once filed, this document would be stored on a justice information system capable of processing XML markup. Then, when a member of the public accesses this document at a courthouse computer terminal or over the Internet, generic text such as a series of Xs—possibly hyperlinked to a message that explains that personal information has been omitted—would be substituted for the tagged personal information. Moreover, when the same document was accessed by the judge or an attorney of record, the justice information system would be programmed to display the entire document, including the sensitive personal information. As this example makes clear, the justice information system would be programmed to respond to a hierarchy of user access privileges, providing to each user the information that he or she is authorized to view. Significantly, such a system would discriminate among users and not the methods by which they accessed the system. A particular user would be granted access to the same information, whether that user was accessing the system at the courthouse or over the Internet.

In this fashion, the next generation of justice information systems will be able to prevent potentially harmful disclosures of personal information, while simultaneously permitting the same degree of public access at the courthouse and over the Internet.

VI. PRIVACY AND PUBLIC ACCESS TO COURT RECORDS OVER THE INTERNET

In reaching its decision in a case, a court will embrace a particular factual narrative of how the dispute being litigated developed. This

107. Id. 22(g)(1) cmt.
narrative is formed from facts found in the case file—the adjudicatory facts viewed by the court as material to its disposition of the case.\textsuperscript{108} Unlike the specific details of sensitive personal information, such as one’s exact street address or the specific numerals forming one’s Social Security number, having access to the adjudicatory facts of a case is essential to achieving the benefits of accountability and citizen education that are promoted by public access to court records. Accordingly, \textit{ceteris paribus}, the public should not be restricted from viewing court documents and files containing adjudicatory facts. In general, courts have accepted this conclusion.\textsuperscript{109}

Recently, however, the argument has been put forward that when adjudicatory facts could prove embarrassing or damaging to one’s reputation, then public access to these facts should only be available at the courthouse. The idea appears to be that making such discrediting or embarrassing adjudicatory facts available to the public over the Internet will result in more members of the public learning of these facts and that this result is not socially desirable—that is, leads to a net loss of social utility. To evaluate this argument, we need to distinguish between discrediting facts and embarrassing facts, and consider in turn the effects of making each kind of information publicly accessible over the Internet.

Discrediting facts are “the sort that impair[reputation and by doing so reduce] one’s opportunities for favorable transactions.”\textsuperscript{110} Discrediting facts often concern “past or present criminal activity or moral conduct at variance with a person’s professed moral standards.”\textsuperscript{111} We might also include within this category, information that “would if revealed correct misapprehensions that the individual is trying to exploit, as when a worker conceals a serious health problem from his employer or a prospective husband conceals his sterility from his fiancée.”\textsuperscript{112} If

\begin{itemize}
  \item \textsuperscript{108} Adjudicatory facts are “unique to individuals—the who, what, where, and when issues typically resolved by juries in judicial trials.” 1 \textsc{Kenneth C. Davis \& Richard J. Pierce, Jr.}, \textit{Administrative Law Treatise} 293 (1994); see also Kenneth C. Davis, \textit{An Approach to Problems of Evidence in the Administrative Process}, 55 \textsc{Harv. L. Rev.} 364, 402 (1942). Professor Davis developed the distinction between adjudicatory facts and legislative facts in the context of administrative agency factfinding.
  \item \textsuperscript{109} Nixon v. Warner Communications, Inc., 435 U.S. 589 (1978) (recognizing a general right to inspect and copy judicial records and documents).
  \item \textsuperscript{110} \textsc{Richard A. Posner}, \textit{Overcoming Law} 539 (1995). They reduce opportunities for favorable transactions because people who will not deal either socially or in business with discreditable people will not deal with a person if they learn discrediting facts about him or her.
  \item \textsuperscript{111} \textsc{Richard A. Posner}, \textit{The Right to Privacy}, 12 \textsc{Ga. L. Rev.} 393, 399 (1978).
  \item \textsuperscript{112} \textit{Id.}
\end{itemize}
Rise of the Machines

discourting facts about an individual are concealed, inefficient social and economic transactions will often result because the decision to transact with that individual will have been made “either with second-rate information or with information obtained at a higher cost.”

Concealing discourting facts about individuals will also lead to an undesirable redistribution of income:

When it becomes more difficult to measure differences among individuals, their treatment becomes more uniform. Lower and higher risk credit are treated as average risk credit, and similarly with the traits of workers, students, and others. It has become a little easier to default on consumer credit, to embezzle funds, and to shirk duties. A redistribution of income takes place within the enlarged class.

Accordingly, as Richard Murphy has concluded, “if accurate information flow is inhibited, there will be an efficiency loss, whether that loss takes the form of increased transaction costs, a cross-subsidization of ‘undesirable’ activity, or simply a decrease in the number of mutually beneficial transactions.” For these reasons, to the extent that public access to court records and documents over the Internet promotes the disclosure of discourting facts, it is efficient, and, a fortiori, restricting such access is inefficient and, therefore, a questionable public policy.


116. Of course, if one believed that informational privacy and the concealment of discourting facts were required by our commitment to values that trumped efficiency, one might very well embrace a different judgment about the merits of such a policy. Today, many privacy advocates of a theoretical bent attempt to justify and ground normative views regarding privacy by appealing to autonomy as such a value. An autonomy-based defense of privacy is generally instrumental, arguing that privacy in one form or another is necessary to protect and further individual autonomy. I find such arguments unconvincing for several reasons that I intend to explore more fully in a separate article. For present purposes, however, three areas of concern may be highlighted. First, whether individual humans are in fact autonomous is deeply problematic and contested. Most autonomy-based privacy theorists simply adopt uncritically an Enlightenment era, modernist conception of the self as autonomous. This conception of the self, however, has been challenged and problematized since the mid-twentieth century by many structuralist, post-structuralist, and postmodern theorists. Accordingly, anchoring one’s defense of privacy in an uncritical and dogmatic notion of human autonomy is to beg an essential question and commit the fallacy of petitio principii. In this regard it is worth noting that even Immanuel Kant did not claim that human beings were autonomous, but merely that an autonomous will is an *a priori* condition on the possibility of moral action—in his sense of “moral action.” See, e.g., Otfried Höffe, *IMMANUEL KANT* 156–57 (1994). Second, given the instrumental character of an autonomy-based defense of privacy, such theorists are promoting a
The preceding argument focuses on static efficiency, that is, the efficient utilization of information that has already been produced. Static efficiency contrasts with dynamic efficiency—the efficiency of the incentives to create new information. While we have just shown that static efficiency favors public access to court records over the Internet, it is also worth noting that dynamic efficiency does not dictate a contrary result. Indeed, assuming that the public has access to court records at the courthouse, the impact of public access to court records over the Internet on dynamic efficiency is likely to be minimal. As a result, the adverse consequences of this enhanced disclosure on the incentives to create new information are likely to be small. Where there is a market for information in court records, data aggregators will compile databases of this information at the courthouse and make them available online for a fee. While this will increase the search costs for obtaining the particular conception of human flourishing and the good life as a socially desirable telos. In a liberal democracy, however, law and social policy should not be grounded in any individual’s or group’s particular conception of human happiness. In the present context, for example, communitarians might very well regard the equation of human happiness with autonomous action as highly problematic and contestable. Finally, insofar as rights to informational privacy appear to require granting an individual some form of control over facts and information, they look suspiciously similar to property rights. In the realm of intellectual property, however, we have a long tradition of viewing facts as part of the public domain. Before we erode this commitment and the public domain that it supports, we should be very sure that there is a sound basis for doing so. The debate over creating a new set of property rights in facts and information should take place within the framework of intellectual property law where if we recognize a property right in facts, we do so directly and expressly after balancing all of the relevant interests and considering all of the relevant policies. We should not do so indirectly through the law of privacy where contested metaphysical appeals to autonomous human selves are likely to be mistaken for an adequate normative and descriptive basis to support such a change.

information, those who value the information more than the access fee will still acquire it. Creditors might be thought a natural example of a group who might pay such an access fee, but they already have access to an inexpensive source for relevant discrediting information: credit bureaus. A more likely example is a passionate gossip, but such a person would probably be just as willing to visit the courthouse. In other words, so long as the information is available at the courthouse for individuals and data aggregators, there will be no shortage of discrediting information available to the interested public. Thus, like considerations of static efficiency, considerations of dynamic efficiency provide no reason for restricting public access to court documents over the Internet.

Embarrassing facts are “the sort that cause[] embarrassment by revealing aspects of a person that while not necessarily or even typically discreditable are not part of one’s constructed public self.”117 While different people will find different things embarrassing, most people desire to keep embarrassing facts about themselves concealed. Accordingly, we may assume that for most people, disclosing an embarrassing fact about themselves is personally undesirable. Whether, however, such a disclosure is also socially undesirable will depend on whether we desire to encourage, discourage, or are indifferent to the embarrassing conduct.

To demonstrate how the socially desirable policy varies with our attitude toward the embarrassing conduct, imagine an individual named Bob who has engaged in some form of embarrassing conduct that resulted in an injury caused by Mary. Imagine further that the circumstances surrounding this injury are such that if Bob brought a lawsuit against Mary, Bob would be awarded compensatory damages. Finally, assume that Bob is rational and has a strong preference for concealing embarrassing facts about himself. Although injured, Bob will not bring a lawsuit against Mary if the expected disutility of disclosing his embarrassing conduct is greater than the expected utility of

117. POSNER, supra note 110, at 539.
compensation for his injury. Now consider the following three cases.

Case 1. Bob’s embarrassing conduct is the kind of conduct that we want to discourage. In this case, the socially desirable policy is to make court records publicly accessible at the courthouse and over the Internet. To see why, consider the effect of this policy both before and after Bob’s decision to engage in the embarrassing conduct. Ex ante, Bob’s knowledge that the court records in any litigation arising from the embarrassing conduct will be publicly accessible will reduce the expected utility of this conduct and (at the margin) will lead to a different, more socially desirable course of action. Ex post, a policy of disclosing embarrassing conduct will punish Bob for engaging in such socially undesirable conduct, possibly deterring Bob and others from doing so in the future. If Bob sues, the punishment will be the disclosure of the embarrassing conduct. If Bob does not sue, the punishment will be the loss of compensation for his injury.

Case 2. Bob’s embarrassing conduct is the kind of conduct to which we are indifferent. In this case, the socially desirable policy is also to make court records publicly accessible at the courthouse and over the Internet. To see why, recall that public access to court records promotes the twin benefits of accountability and citizen education. Because we are indifferent to Bob’s embarrassing conduct, whether Bob engages in this embarrassing conduct will neither increase nor decrease the social utility derived from these twin benefits. Moreover, since Bob is rational, he will act to maximize his personal expected utility. Thus, the policy of making court records publicly accessible enhances social utility to the extent that it includes both Bob’s personal happiness and the benefits of accountability and citizen education.

Case 3. Bob’s embarrassing conduct is the kind of conduct that we want to encourage. In this case, the socially desirable policy is to limit public access to court records at the courthouse and over the Internet to the extent necessary—and only to the extent necessary—to avoid the

---

118. The *expected disutility of disclosure* is equal to the probability of disclosure times the disutility of disclosure. The *expected utility of compensation* equals the probability of winning the lawsuit times the utility of the damages. Under a legal regime in which court records are private or sealed, the probability of disclosure approaches 0. Under a legal regime in which court records are public and easily accessible, the probability of disclosure approaches 1.

119. The three cases set out below assume that Bob is the agent of the embarrassing conduct. Cases could also be developed in which Bob is the recipient or victim of conduct that he finds embarrassing to report. The same policy recommendations would result.
embarrassment of disclosure. Under this policy, Bob does not need to fear the disclosure of his embarrassing conduct. As a result, he will bring the lawsuit against Mary and receive compensation of his injury. This policy encourages the embarrassing conduct both ex ante and ex post. Ex ante, it encourages the conduct because its expected utility is not reduced by the prospect of disclosure. Ex post, it encourages it because he is not punished.

Under this third scenario, limiting public access to court records at the courthouse and over the Internet to the extent necessary to avoid the embarrassment of disclosure can be achieved in two ways. First, for embarrassing conduct that is of a type that recurs frequently, an exception to the public access rule can be established for cases that often involve this kind of conduct. This approach is already used to encourage victims of conduct that they might find embarrassing to report to seek the protection of the courts. Exceptions would presumably be made for cases involving juvenile dependency (abuse and neglect), termination of parental rights, petitions for waiver of parental consent for minor abortions, adoption, guardianships, conservatorships, mental health, and sterilization. Second, for embarrassing conduct of a type that is unlikely to recur, courts can be permitted to limit public access to court records and case files on a case-by-case basis. Courts already exercise this power when they entertain motions to seal documents and entire case files.

Whether one limits access by a general exception to the public access rule or on a case-by-case basis, it should only be limited to the extent required to prevent embarrassment to the potential plaintiff. In most cases, this would require only that all personally identifiable information pertaining to the potential plaintiff be removed from the public records. As already noted, this could be done easily if the information in court records is marked up with XML tags.  

Finally, we need to consider the impact of making court records containing facts that are both discrediting and embarrassing publicly accessible over the Internet. As before, whether the disclosure of such facts is socially desirable will depend on whether we desire to encourage, discourage, or are indifferent to the underlying conduct. If we want to discourage or are indifferent to the embarrassing and discrediting conduct, then insofar as revealing discrediting information is efficient, the fact that it is discrediting in addition to embarrassing

120. See supra Part V.
121. See supra notes 110–15 and accompanying text.
simply strengthens the case for public access over the Internet and at the courthouse that was made above for merely embarrassing conduct.  

Since presumably we would never want to encourage any kind of discrediting conduct, the only interesting case remaining is one in which we want to encourage the victim of discrediting and embarrassing conduct to seek the protection of the courts. Consider, for example, a woman who while having an abortion is injured through the doctor’s negligence. As noted in our earlier discussion of merely embarrassing conduct, to encourage such victims to seek recourse in the courts, we should restrict public access both at the courthouse and over the Internet to the extent necessary to avoid embarrassment to the victim. In the present example, however, there is a public interest in disclosing the doctor’s malpractice. While the same information cannot both be revealed and concealed, we can nonetheless develop a process that respects both the needs of the victim and the interests of the public. As with cases of merely embarrassing conduct discussed above, if such a case recurs frequently, it should be subject to a categorial exception from the public access rule, and public access to the case file should be restricted both at the courthouse and over the Internet to the extent necessary to avoid embarrassment to the victim. If, however, such a case occurs only rarely, then the victim should be able to petition the courts to restrict public access to the case file at the courthouse and over the Internet to the extent necessary to avoid embarrassment to the victim. Under either approach to restricting public access to the discrediting and embarrassing facts, a court record indicating the defendant’s (e.g., the doctor’s) involvement in a lawsuit should be made publicly accessible both at the courthouse and over the Internet, and a process should be created that permits individuals with a legitimate interest in acquiring the discrediting information to petition the court for access to the restricted information. Thus, like facts that are either discrediting or embarrassing, facts that are both discrediting and embarrassing do not justify treating public access to court records over the Internet differently from public access at the courthouse.

The foregoing discussion suggests that notwithstanding the presence of discrediting and embarrassing adjudicatory facts in court records, documents, and files, the public should be able to access these records

122. See supra Cases 1 & 2.

123. For an example of such a procedure, see STEKETEE & CARLSON, supra note 63, at 53 (CCJ/COSCA Guideline § 4.70(b)).
over the Internet to the same extent that it can access them at the courthouse. If court records and case files should be sealed, then public access to them should be prohibited both at the courthouse and over the Internet. If they are accessible at the courthouse, then they should be accessible over the Internet as well: the courts should not discriminate between methods of access.

VII. CONCLUSION

The machines are coming—a new generation. XML-enabled, these new machines are able to connect to existing legacy systems and with each other to form local, state, and national networks. From the myriad connections being established among the local machines and information systems of courts, justice agencies, law enforcement, correctional facilities, social services and treatment providers, and other members of the justice and public safety communities, integrated justice systems are emerging. These integrated justice information systems provide significant efficiencies, among them diverse kinds of cost savings, error reduction, and improvements in productivity. They permit members of the justice and public safety communities to seamlessly interoperate and exchange records, files, and messages, bridging the information chasm that once separated islands of isolated data. In light of these advantages, the continuing integration of justice information systems into ever larger systems appears inevitable. The Office of Justice Programs at the United States Department of Justice and the Organization for the Advancement of Structured Information Standards are developing XML standards to further facilitate this integration. Court administrators and policymakers mindful of the benefits within their reach are implementing XML solutions.

XML solutions enable individual courts to integrate their case management information systems with an electronic docket and their document management system. They permit the courts to migrate from an increasingly anachronistic paper-based conception of a case file as a collection of documents and records to an understanding of it as structured information. With structured information, one is no longer limited to manipulating two-dimensional containers of information, such as a sheet of paper or the surface of an LCD, but can process and manipulate the information itself. This capability by itself promises to revolutionize court work flow.
These same XML technologies enable integrated justice information systems to interoperate and exchange data over the Internet with personal computers owned by members of the public. With this new ability arises a new question of policy: whether public access to court records over the Internet ought to be permitted and, if so, to what extent? There are at least three answers to this question: public access to court records over the Internet should be denied; public access to court records over the Internet ought to be more limited than the access available at the courthouse; and public access to court files and documents over the Internet should be permitted to the same extent that the public has access to them in the courthouse—that is, courts should not discriminate with respect to methods of access. As the readers who have perused the preceding pages know, I believe that the last is the better answer.

To defend this position, I have parsed the information in a case file into two general categories: sensitive personal information and narrative adjudicatory facts. With regard to the former, I argue that for reasons of public safety and personal security, the public should not have access to it either at the courthouse or over the Internet. Accordingly, the presence of such information in the case file does not provide any grounds for treating public access to court records over the Internet differently from public access at the courthouse. Along the way, I discuss two different ways of removing sensitive personal information from the public case file, one using paper or electronic forms, and the other using XML. Unsurprisingly, I believe that courts will ultimately adopt the latter solution—even if they are forced to use forms in the short term as a transitional measure.

With regard to narrative adjudicatory facts, I further subdivide this category into discrediting facts and embarrassing facts, implicitly recognizing a third subcategory containing all others. Among narrative adjudicatory facts, discrediting and embarrassing facts are foregrounded because they provide the most compelling case for restricting public access to court records over the Internet while allowing unrestricted public access at the courthouse. Clearly, such facts would be more easily and more widely known if the public was permitted to access them in court records over the Internet. Nonetheless, I argue that neither discrediting facts nor embarrassing facts, nor indeed facts that are both discrediting and embarrassing, provide sufficient grounds for limiting public access to court records over the Internet while permitting unlimited access to them at the courthouse. As to discrediting facts, I make my argument by appealing to considerations of static and dynamic
Rise of the Machines

efficiency. As to embarrassing facts, both benign and discrediting, my argument proceeds inductively: considering the effects of disclosing embarrassing facts under several different scenarios. While I conclude that embarrassing facts do not justify discriminating between methods of public access, I do note that under some circumstances, the public should be barred from accessing embarrassing facts both at the courthouse and over the Internet. I note as well that when appropriate, XML permits us to easily remove embarrassing facts from the public files.

The machines are coming. XML-enabled, they promise a justice system that is more efficient, accessible, and humane.
DESIGNING AN ACCESSIBLE, TECHNOLOGY-DRIVEN JUSTICE SYSTEM: AN EXERCISE IN TESTING THE ACCESS TO JUSTICE TECHNOLOGY BILL OF RIGHTS

T.W. Small,* Robert Boiko† & Richard Zorza‡

Abstract: The Access to Justice Technology Bill of Rights project, sponsored by the Access to Justice Board of Washington State, included a committee composed of attorneys, judges, technologists, and librarians charged with envisioning an ideal civil justice system. Our goals were to design a system with certain core values (e.g., due process and access to justice), test the system using a complex family law scenario, determine what opportunities technology brings to the table, and identify what barriers technology creates for persons using the system. This Article describes an idealized civil justice system (System) unlike anything that presently exists. The System is composed of people and technology that together provide a factual information-delivery system, an advocate, an adversary, a mediator, an adjudicator, and a proactive enforcer. To be successful, our System needs to use a wide variety of current and next-generation technologies and processes. The System gives the participants in a legal issue the opportunity to resolve their issue by themselves before escalating the issue for mediation or adjudication. In addition, the System plays an active role in the enforcement of whatever resolution is reached. At the core of the System is a cycle in which all participants simultaneously review and choose options. The interaction of all the participants choosing options allows the System to converge to a mutually acceptable resolution of the issue.

I. INTRODUCTION

What if twenty years from now you were founding a new country and you were assigned the task of designing its first civil justice system? What if you were not limited to developing a traditional adversarial justice system? Suppose the system could use any technology that is currently available but had to maintain these core values:

* Judge T.W. “Chip” Small, Superior Court Judge for Chelan County, Washington; J.D., University of Washington; past chair of the Access to Justice Board for the State of Washington; former chair of the Access to Justice Tech Committee; and co-chaired the Opportunities, Barriers and Technology subcommittee for the Technology Bill of Rights project and received the ATJ Community’s Judicial Leadership Award in 2001.
† Bob Boiko, Associate Chair of the Masters of Science in Information Management at the University of Washington’s Information School. President of Metatorial Services Incorporated, a Seattle-based information systems consulting firm and author of the Content Management Bible published by Wiley.
‡ Richard Zorza, Esq., A.B., Harvard University, 1971, J.D., Harvard University, 1980, has been the primary consultant to the Access to Justice Technology Bill of Rights process.
Availablity of the system to all people regardless of their income level, prior experience, education level, and cultural background;

Fairness of the system to all participants;

Confidentiality of the system and protection of participants’ privacy rights;

Ability of the system to reach a resolution that maximizes the interests of all participants; and

Openness of the system to scrutiny by the public.

What would your justice system look like? The Opportunities, Barriers and Technology Subcommittee of the Access to Justice Board, composed of knowledgeable technologists, attorneys, librarians, and judges, explored such a system. Our goal was to design a system, then test it using a complex family law scenario. In our work, we sought to determine the following:

What process our proposed system might use to elicit just resolutions of civil issues?

What opportunities technology might offer?

What barriers to justice the system might create for the people who use it?

The new civil justice system (System) we designed is unlike anything that presently exists because of the iterative and escalating process it uses. In the System we envisioned, participants are in a continual cycle of review and decision. They review the background of their case and the options that are currently open to them, and then decide on what their next step might be. If two participants can resolve their issues in this way without intervention, they have no need of lawyers and judges. If they cannot resolve their issues in this way, the System provides the ability to escalate the issue through a mediator and then finally through an adjudicator. In addition to providing this escalation, the System serves alternatively as a neutral information-delivery system, advocate, adversary, unbiased decision-maker, and proactive enforcer.

This Article describes the new System we envisioned in the context of a family law scenario to be resolved within the System. We discuss the issues raised by the use of the System’s technology and describe the base technologies envisioned. We then describe the opportunities and barriers inherent in the System and present suggestions for avoiding or minimizing them.
Testing the Access to Justice Technology Bill of Rights

The Access to Justice Technology Bill of Rights Process is designed to do more than come up with a set of rules for technology in the justice system—it seeks rather to suffuse the justice system with a set of values that will guarantee that technology will always be used to increase rather than limit access. Perhaps the greatest challenge is to find a way that will stand the test of the long term, take advantage of all the opportunities, and be a tool to resist all the dangers presented by the inevitable multiple and unforeseeable changes in the technological environment.

A multi-disciplinary committee of judges, lawyers, technologists, and lay people worked throughout 2002 and 2003 to create a test bed for this endeavor. Their idea envisioned a civil justice system that took advantage of every technological opportunity and refused to be constrained by the status quo or current processes. The fundamental requirements of due process, access, and the realistic potential of foreseeable technology were the only real constraints. It is hoped their product will illustrate the potential of technology to empower access; provide a yardstick against which innovation can be judged; and highlight the areas of danger in the deployment of technology, areas that the Access to Justice Technology Bill of Rights must be strong enough and broad enough to protect against. The group articulated the fundamental requirements of due process as “an easy path to interactions where a person’s legal concerns, needs and/or rights are properly presented, listened to, given meaning and dealt with impartially and carefully.”

We assumed the lawmaking body would be separate from the civil dispute resolution system, and that all changes in substantive and procedural laws and rules would be immediately incorporated into the system. Furthermore, we anticipated that the lawmaking body would have access to the System and the results it produced. Legislative decisions would be enhanced by the feedback the System would provide.

II. OUR SCENARIO: THE RODRIGUEZ FAMILY

Mr. and Mrs. Rodriguez were married seven years ago. Mrs. Guadalupe (Lupe) Rodriguez is a monolingual, literate Spanish speaker who now seeks divorce from her husband, Eduardo. Eduardo is illiterate in Spanish and English, but speaks both languages. The couple’s children, Maria and Diego, have been living with their mother since the

---

1. The names used are fictitious. Any resemblance to a real family is unintentional.
couple separated eight months ago. Eduardo has spent a great deal of
time with the children, before and after the separation.

Financially, the couple is struggling. Eduardo works full time and
Lupe is on public assistance. The couple has significant credit card debt
and is two months behind on the rent on the family home where Lupe
and the children still reside. The landlord has threatened eviction.

Both parties want custody of the children, child support, and the
family home. Lupe is seeking payment of outstanding bills. Also, Eduardo has accused Lupe’s brother-in-law of abusing his children.

The participants in this issue include:

Eduardo and Lupe Rodriguez;

The Department of Social and Health Services, which provides
public assistance and may investigate the claim of child abuse;

A guardian ad litem assigned to advocate for the best interests of
the children;

The children’s school, which was notified of the couple’s
pending divorce and custody dispute;

The credit card companies to whom the couple owes money;

The landlord to whom the couple also owes money; and

Community interests, such as the media and the general public,
who, respectively, may seek access to public information and
accountability of the System.

Each of these participants has some stake in the issue posed and so
has some right to information regarding it and a right to some input into
its eventual resolution.

III. ASSUMPTIONS BEHIND THE SYSTEM

With this scenario and key players in mind, the subcommittee
envisioned a System to help the parties resolve their problems in the
most just and efficient manner possible. The following subparts detail
the initial constraints that we put on this new System. In forming these
constraints we sought to look beyond the current structure of the United
States legal system to find instead the motivations behind its concept of
justice.
Testing the Access to Justice Technology Bill of Rights

A. The Core Values

There are basic core values supporting any system of justice. For our System, we sought to design an easy path to interactions where a person’s legal concerns, needs, and rights are properly presented, listened to, given meaning, and dealt with impartially and carefully. Thus, in keeping with its charge, the committee assumed the System must provide opportunities without presenting barriers to access.

We agreed that the System must employ the best balance between people and technology. Where judgment is concerned, it must employ qualified people. Where information processing, record keeping, communication, or other important but non-cognitive processes are concerned, the System ought to rely on technology. Thus, the proper system augments human judgment with sophisticated automation so that judgment may be applied as effectively as possible.

B. System Roles

We recognized the following types of roles or personas that the System would need to fill in order to completely serve the needs of the participants in any issue:

Neutral Information-Delivery System. At many points in the resolution of an issue, participants simply need information without any added perspective, opinion, or value judgment. For example, Eduardo may simply need to know the total amount of back rent that the landlord feels is owed.

Advocate. At some points, the System must advocate for the needs of a particular participant. For example, the System should help Lupe construct her arguments in favor of her getting custody of her children. The System is not intended to replace human advocates. If Lupe has a lawyer, the System should, in fact, facilitate their interactions.

Adversary. When presenting the views of participants that are in disagreement, the System may appear to take an adversarial position. For example, the System may report, against Lupe’s wishes, the alleged abuse by her brother-in-law.

Unbiased Decision-Maker. In presenting options and in any mediated or adjudicated decision, the System must take on the role of a disinterested authority. For example, if the parties reach
adjudication and custody is awarded to Eduardo, then the System must deliver this decision with unbiased authority.

*Proactive Enforcer.* Once a resolution has been reached, the System must help uphold that resolution. For example, if custody is awarded to Lupe, the System must then appear to Eduardo as a law enforcement official.

The committee recognized that participants must at all times be aware of what role the System is playing and that those roles must never become confused in the participants’ minds.

C. The Idea of an Issue

We devised the idea of an issue to cover the wide range of scenarios that the System may need to work through. An issue is very likely to be a complex set of interrelated sub-issues. For example, the summary issue in this scenario is the custody contention between Eduardo and Lupe. Clearly that is not a single issue, but rather a congregation of issues including, but not limited to, child-abuse, rent collection, and school attendance. Each of these issues, in turn, may be composed of its own set of sub-issues. The System must be capable of tracking and interrelating these sets of issues so that as each is resolved, its solution feeds into and helps to guide the remaining issues toward resolution. For example, if the allegations of child abuse are shown to be valid, then the impact of that resolution will need to be felt throughout the rest of the System. Options that were once available may no longer be and new options for resolution of other issues may appear.

Consequently, participants are linked together in the System by being parties to an issue. Issues can have any number of sub-issues, each with its own set of participants. Resolution of the top-level issue is contingent upon resolution of its sub-issues. When all the sub-issues are resolved, the main issue that originally brought the participants to the System can be resolved as well.

IV. THE SYSTEM AND ITS STAGES

The core solution is a four-stage process. While it resembles the current legal system, it is actually very different because of the technology used to manage the interactions within each step, the way decisions are made within each step, and the manner in which the case moves from one step to another. These stages will be explained in detail using the hypothetical scenario.
Testing the Access to Justice Technology Bill of Rights

Participants need to move from initial contention to final coordination on the resolution of their issues. The System has a series of possible stages that take participants from their initial contact with the System to final resolution of their issues:

- **Start-up**, where participants first make contact with the System;
- **Interaction**, where the participants try to resolve their issues without outside intervention;
- **Intervention**, where the participants are helped, more and more proactively to reach resolution if they cannot reach it alone; and
- **Compliance**, where the resolution is put into effect and the participants adhere to it.

### A. Stage One: Start-Up

For the System to be of value, each participant must be aware of its existence, have relatively easy access to it, and have assistance available when necessary. How do Lupe and Eduardo know that the System is available to resolve their problems? How do they interact with the System given their language and literacy barriers? In particular, we concluded that these factors are essential to the start-up process:

- **Awareness**, how members of the public are made aware of the existence, purpose, and functions of the System;
- **Outlets**, how persons can contact the System;
- **Intermediaries**, persons who assist individuals unwilling or unable to use the technology; and
- **Initiation**, how interaction is begun between an individual and the System.

#### 1. Awareness

Public awareness of the System must be pervasive. Publicity of the System might involve:

- Television commercials, which present typical legal problems and an example of successful resolution using the System;
- Billboards in multiple languages advertising the location of points of access to the System; and
- Outreach to community leaders including religious, educational, political, and social leaders to enable them to make their constituents aware of the availability of the System. This
outreach may include appearances by people who have successfully resolved issues using the System. The message delivered should reinforce the core values of the System. In our scenario, we assume that Lupe first heard about the System from a friend at church. An intermediary gave Eduardo notice that Lupe initiated an issue and advised him of how to interact with the System.

2. Outlets

The System is designed to be accessible to anyone. Fundamentally, that means that the System cannot be a typical computer program that requires special knowledge to be used. What differentiates those who use the System is their level of comfort and direct interaction with the System. Some people will be able to access the System directly and take personal control of their interaction with it. These individuals will be able to access the System through any networked computer without the need for assistance from intermediaries.

Others, like Eduardo and Lupe, will need or want an intermediary to assist them to work through the System. These participants will access the System in a public (e.g., public library) or private (e.g., church) outlet where a trained System intermediary will guide them through the System. Thus, regardless of a person’s training or experience, the System will be fully available to them.

3. Intermediaries

Three types of people stand between participants and the System. The first is an appointed agent of the System. These people are empowered to officially represent the System to participants. For example, in the case where notice needs to be served to Eduardo (who is illiterate), an agent would be appointed to deliver it verbally and to advise Eduardo of his options. Agents provide the most accessible level of interaction with the System. They are multilingual and trained to effectively communicate a person’s legal responsibility to interact with the System and the options that person currently has for interaction.

Agents are officials of the System. The other two types of intermediaries are not. Information professionals help people contact the System and communicate with it. Attorneys and other legal professionals provide advice on what options participants ought to pursue.

Information professionals guide the participant to the proper place in the System and leave it to the individual to decide what action to take.
Testing the Access to Justice Technology Bill of Rights

They do not interpret the information the System provides or give advice on what are the best options.

We imagine that information professionals (hired by the state or by public access organizations) will be located in a variety of locations including:

Government buildings such as libraries, post offices, and courthouses;

Community centers, churches, and other civic buildings; and

Public legal services offices where navigation services are provided by intermediaries. Attorneys may also be available to provide advice.

We foresee that private System intermediaries will be available to provide fee-based services. Private System intermediary offices will offer premium service and value added features, such as social service referrals and additional proprietary information to form the basis for charging a fee. Private intermediaries may also be employed by larger organizations to assist their employees.

Legal professionals may serve as information professionals, but their focus is on advice rather than access. We envision, for example, that law firms would provide both access to the System and advice on what options to pursue. So, in the System we envision, legal professionals provide much the same service as they do today. However, their services would be centered more fully on advice about the various options presented to a participant because much of the organizational and administrative interactions would be handled by information professionals.

4. Initiation

We envision two ways a person may initiate interaction with the System: (1) a non-official mode where the individual merely wants to experiment with certain scenarios and (2) an official mode where the individual wants to initiate use of the System. The official mode will also include those persons drawn into interaction with the System by action initiated by someone other than themselves.

Non-official mode is a way that individuals may learn how to use the System and become comfortable with it. Those individuals who want to understand how an issue may play out, get familiar with an issue, or just learn about the System, may use the System in non-official mode. In this mode, all information is anonymous. They may also familiarize
themselves with how the issue that they are actually immersed in may be resolved in order to assist them in forming their own personal strategies and to chart probable outcomes. The individual may set up the issue in a variety of ways and track the ramifications of certain actions on his or her part. The individual can then “undo” options and attempt to play out the issue in another manner. The non-official mode allows the user to involve as many players as desired.

Official mode is used when a person wants to initiate action on a real issue or has been drawn into the System by being named in an issue by someone else. In official mode all participants are required to be identified. However, their identity will only be revealed to other participants on a need-to-know basis. Only the minimum amount of personal information necessary for a participant to exercise her or his options within the System will be available. Once options have been exercised there may be binding consequences to that participant. Certain choices will have time limits with penalties for not responding within the prescribed period of time. Participants will be bound to follow the process through completion.

The church’s computer put Lupe in contact with the System, and she spoke in Spanish to an information professional who guided her through the initial process. After exploring her options with the intermediary, she chose to begin the divorce proceeding. A System agent contacted Eduardo at work and gave him notice of Lupe’s action and explained his options.

B. Stage Two: Interaction

Within the interaction stage, the System we envisioned uses a repeated cycle to move participants as quickly as possible toward resolution of their issue. Once initiated, an issue enters the spiral of resolution—notice, exploration of options, exercising options, and feedback:

Notice, where participants are told the current status of their issue and given a set of options to move the issue forward;

Exploration of Options, where participants educate themselves on the options, play out the possible results of their options, and decide what option to take;

Exercising Options, where the participants choose an option; and
Testing the Access to Justice Technology Bill of Rights

Feedback, where participants are told of the options exercised by other participants and what that means to them. Feedback leads to further notification when new options become available due to the choices participants have made. This process is illustrated in the following diagram:

The System is spiral in nature. As participants get notice of an action, they explore and then exercise an option. After they act, other participants receive notice, explore, and exercise further options in response. If all goes well, the process converges on a resolution that represents the best fit between the desires of all parties. If the participants reach an impasse, the System escalates to a more proactive stance (mediation and then adjudication) to move the issue forward towards final resolution.

The success of this methodology rests on the System’s ability to determine all possible options at any point in an issue and to effectively present them to each participant. We believe that it is not feasible to
begin by creating some enormous logic of options that covers every possible legal issue. Rather, we believe that the System would begin with a limited logic that augments the judgment of options of a trained judge or other official of the System. However, if the System is built to consistently learn from these officials, then over time the logic of options will become more and more robust. In the long run, only those situations that are truly new or extremely complex will require human judgment. Similarly, the effective presentation of options and the background information needed to explain them will grow as the System learns from and reiterates the human decisions made over time.

1. Notice

Notice is the way the System communicates options and background information. Notice:

- Has basic information about the issue, the initial options available, and the resources available to understand and evaluate the options;
- Is sent in different modes depending upon the needs and abilities of the recipients of the notice (e.g., electronically, on paper, verbally, in English, or other languages);
- Is sent to all participants. Participants are those people who can exercise options. Participants may be named in the issue (for example, Lupe names Eduardo), or unnamed (for example, the landlord has collection options that can be exercised even though not specifically named by Lupe); and
- Is sent to interested parties. Interested parties do not exercise options but still need to know about the issue. For example, a notice might be provided to the children’s school, so that their teachers are aware of their situation.

2. Exploration of Options

After notice is issued, participants may then start to explore their options. The options are those provided by the System. The System will provide current and relevant information to assist the participants in selecting a particular course of action. To explore options, participants may:

- Study the background information described in the notice. Background information can be delivered in the same modalities
Testing the Access to Justice Technology Bill of Rights

as the notice itself (e.g., digitally, in print, or verbally in a variety of languages); and

Move into non-official mode to play out certain options. As discussed above, in non-official mode, participants can experiment with various options and see what the possible outcomes might be. The participants can run and store these simulations for later access.

In our scenario, Eduardo’s first notice gives him information about the impact the divorce will have on his children. While experimenting with his option to give Lupe custody, he calculates the likely amount of child support he will have to pay. He decides not to contest the divorce, but to seek custody of the children. He also learns of the alleged abuse of the children by Lupe’s brother-in-law and learns how to file a complaint with the state’s Child Protective Services (CPS).

3. Exercising Options

After participants have researched their options, they must exercise one. As with notices, they can respond in any mode. When the last option is exercised, the System moves into the compliance stage. In our scenario, Eduardo chooses to seek custody and file a CPS complaint; and the cycle of notification and exercising options will continue.

During the choice process, the System’s main goal is a convergence of the options selected by all participants. In our scenario, a trivial convergence would occur if Eduardo assented to Lupe’s request for divorce and custody. Of course, such easy resolution can hardly be expected. Thus, the logic of options the System employs has the tough job of gently pulling the participants toward a convergence of interests without introducing any bias toward a particular outcome. Rather, the System seeks a “best fit or equilibrium” between the different actions of the participants.

In addition, the committee recognized these as key aspects of the System’s behavior during the choice process:

*No preset time constraints.* The System has the capacity to wait for responses indefinitely. However, participants do not always have an unlimited amount of time to respond or arrive at a final resolution. The timeframe for participants to respond is determined by rules of civil procedure and applicable statutes. For example, Eduardo is advised that if he does not respond to Lupe’s action within a certain timeframe, then Lupe will prevail.
automatically. In addition, CPS is required to investigate Eduardo’s complaint within a set amount of time.

*Continuous education.* The educational aspect of the System is continuous and remains active throughout the process. Whenever a participant has questions about procedure, substantive law, or the likely outcome if a particular option is chosen, the System is there to inform and educate the participant.

*Complete issue histories.* Throughout the process, the System logs responses and any other information that has been provided by the participants. Within privacy and security constraints, participants can always find and review this information.

*Parallel processing.* Participants simultaneously interact with the System. Unless the options of one participant are dependant on those of another, the System allows people to work at their own speed.

4. *Feedback*

The option selected by one participant becomes a source of feedback to others in the System. The participants affected are then allowed to explore their options and, in turn, exercise them to provide feedback to the other affected participants. In our example, Lupe, the school, and CPS are notified of Eduardo’s decision to file an abuse complaint. As with e-mail, and other modes of asynchronous communication, each participant receives notice, acts, and awaits the next communication. All participants simultaneously receive notice, explore options relevant to them, and make choices that produce feedback for others in the System.

Simultaneously the System monitors the responses and checks to see whether convergence is occurring. If the System detects a convergence, it can then propose a resolution to the participants. If they agree, then they have achieved equilibrium and the System moves on to the compliance stage. If the System detects an impasse or deadlock, then intervention will be necessary. This process is illustrated by the diagram below:
Testing the Access to Justice Technology Bill of Rights

In our scenario, after Lupe finds out the results of the CPS investigation, she may realize that Eduardo’s behavior is justified. After both of them learn more about the impact of divorce on their children, they may take the option to reconcile and terminate the process. On the other hand, Lupe may decide it is in her children’s best interests to continue to seek a divorce and custody of her children. If Eduardo continues to believe it to be better for the children to live with him, then an impasse has occurred, and intervention will be needed to move the process forward.

C. Stage Three: Intervention

The System is designed to provide all participants the information and options necessary for them to reach their own resolution. However, if the participants’ positions are clearly defined, and the information necessary for a resolution has been presented without an agreed resolution, then the process is at an impasse. At that point, we envision the System breaking the impasse by either mediation or adjudication.
1. Mediative Role

If the participants cannot agree by themselves, then the System takes on the role of neutral mediator. As in all parts of the System, the mediator is a human augmented by an automated data processing and delivery system. The mediation process is governed by the following principles:

Questions on need. The System asks questions of the participants to find out what the participants really want or need. It gives them feedback and answers based on the answers to those questions.

Suggestions based on history. The System suggests prior acceptable agreements in other factually similar situations as ways to resolve differences.

Suggest terms. The System suggests prior terms that worked in other similar situations.

Advise about consequences. The System tells participants the consequences of non-agreement in other similar situations.

Information disclosure. When participants are given recommendations by the System, the data that the System’s projections are based upon is shared with the participants.

Discretion. The human mediative aspect of the System allows discretion of the mediator to control whether to give data on outcomes to individual participants. In exercising discretion, the System will weigh the risks of incentive not to settle to determine whether releasing the information is likely to promote resolution or impasse.

Restriction. Part of the System’s information will not be accessible to all participants. The System gathers some information that is used solely for mediation purposes.

Fair distribution of information. The System is designed to ensure that all participants have access to the same information and to complete information from the relevant sources.

Equal bargaining ability. The System prevents more forceful and capable participants from taking advantage of less capable participants, at an economical and emotional cost to the participants and others (particularly children).

In addition, voluntary resolution of problems is subject to the approval of the System whenever there is unequal bargaining power or
Testing the Access to Justice Technology Bill of Rights

vulnerable third parties (such as children). This review is intended to prevent coercive and unfair resolutions. Thus, the mediative role of the System may act to ameliorate problems vexing the litigants as opposed to only being an impartial fact finder and independent decision maker.

In our scenario, the System uses a wide range of electronically gathered and analyzed information such as credit reports, bank records, and other financial information. In addition, information from the school system, CPS, and other state agencies is made available to Eduardo and Lupe to help them resolve their issues. Also, with the consent of the parties, hidden video can be used to assist the guardian ad litem to observe the interaction between the participants and their children. Finally, the System uses a database of prior issues with similar constraints to show the participants the most likely outcomes. Together, these tools give the System the authority and the information to help break the impasse. To play the System out to its conclusion, however, we assume that Eduardo and Lupe still cannot agree to a solution and an adjudicator must intervene.

2. Adjudicative Role

The System must be able to assume a more traditional adjudicative role. Despite the best efforts of the System as mediator, our committee recognized that the System’s efforts will not always be successful. The System would use data of past decisions to determine the reasonableness of the current decisions and their impact on the participants. Like mediation, adjudication is a combination of human and automated processes.

Many financial issues may be automated effectively. When the objective evidence dictates a decision with a very high degree of certainty (such as child support) then the likeliest decision may be arrived at automatically. The legal ownership of individual property and the fair market value of the property, for example, can often be resolved through the use of technology or the information gathering that occurs prior to the participants reaching impasse. Furthermore, the System could automatically update financial orders such as child support, so that participants need not return to the System for routine updates. Even at an impasse, minor issues can be redirected to adjudication or mediation to ensure that the adjudicator only deals with those human issues that cannot be resolved through the use of technology or voluntary resolution.
Adjudicators can depend on the information already gathered by the System from all participants to considerably ease the fact finding that they do in the current legal system. In addition, a report could be provided by the System to the participants so they can see the consequences each factual finding has on the remainder of the process.

We envisioned that the System’s ability to gather and verify information will have major impacts on the reliability, efficiency, and long-term stability of the System. System-gathered data is treated as presumptive rather than conclusive, with the strength of the presumption depending upon the nature and source of the information. In the System we envisioned, the role of advocates changes from mainly information gatherers to presenters who argue the implication of the information by reducing the need for discovery during the adversarial process. The System automatically resolves some issues and highlights those that need a human adjudicator. Advocates use the greater reliability of information that has been gathered by the System to become more effective. Discovery abuse is minimized and fairness is enhanced.

The adversary process that ultimately resolves issues that cannot otherwise be resolved is enhanced by de-emphasizing discovery. Furthermore, in the advocate role, the System can use technology to provide legal advice by remote access to court hearings by pro bono counsel for those participants who the mediative aspect of the System determines cannot make effective presentations during the adjudicative phase. The System can also allow participants to see the probable consequences of agreeing to settle at any point during the adversary process. The adjudicator will also have statistical information available on long term outcomes associated with each possible judgment. The System will strive to have the adjudicator impose a resolution in as few a number of issues as necessary, seeking voluntary resolution of most issues.

In our scenario, the System assigns an adjudicator to the issues facing Lupe and Eduardo. That person reviews the trail of events that have lead to the impasse. The adjudicator focuses on those issues that might break the impasse and allow the two parties to still reach agreement. Each of the parties has additional opportunities to argue their case to the adjudicator. After taking the previous interactions and the more recent arguments into account, the adjudicator decides to grant a divorce to the couple and award custody of the children to Eduardo. The adjudicator then sends the two parties back to the mediator to work through the issues of visitation and child support.
Testing the Access to Justice Technology Bill of Rights

D. Stage Four: Compliance

Today’s courts are not proactive enforcers of their decisions. Ensuring compliance today is dependent upon the prevailing party bringing non-compliance to the attention of the court before the court may take steps to ensure compliance. Other than automatic liens on real property in the county in which a monetary judgment is rendered, the courts do not actively enforce compliance; rather, they reactively enforce compliance.

The justice system envisioned by our committee involves a System that is proactive and self-executing when it comes to compliance with most resolutions, regardless of whether the resolutions are voluntary. Some of the tools we envisioned the System using to proactively ensure compliance were:

- **Reporting requirements.** Participants against whom orders are imposed are required to report their own compliance or non-compliance to the System. Failure to report compliance creates a presumption of non-compliance. Information about non-compliance is then provided to the participants in whose favor compliance was ordered. Those participants are then asked whether they want compliance enforced. If they do, the non-complying participants are given the ability to rebut the presumed non-compliance after receiving notice using the same process described to resolve problems. It is at that time the non-complying participants are given the opportunity to explain.

- **Detailed order.** Orders are detailed enough to inform participants of the consequences of non-compliance. For example, consequences of failure to exercise visitation rights may result in suspension of those visitation rights.

- **Reminders.** The System reminds participants of the need to comply with its orders.

- **Resources.** The System provides information and resources to enable compliance.

- **Voluntary modifications.** The System allows the participants to voluntarily agree to modify compliance when necessary.

- **Outcome correlations.** The System informs the adjudicator and the participants of correlations between early non-compliance and long-term consequences on the participants and third parties.
Information sharing. The System shares information about non-compliance with other government agencies.

The System we envisioned is more proactive in obtaining compliance of its orders and continues to educate the participants of the consequences of non-compliance to promote voluntary compliance with the System’s orders.

In our scenario, the System advises the concerned third parties (the children’s school, social services, media, etc.) of the parts of the decision that they are entitled to know. In addition, all of the forms necessary to finalize the divorce and custody issues are automatically assembled and delivered to the parties. Finally, the two parties are given a final notice that explains their options for modifying or appealing the decisions that have been made as well as a list of resources (information sources, public and private agencies, and so on) that can help them understand and comply with the decisions that have been made.

V. THE TECHNOLOGY BEHIND THE SYSTEM

The Opportunities, Barriers and Technology Subcommittee brought together legal and technology experts to define the set of current and next generation technologies that might be part of the System we were envisioning. Those technologies are listed below. Because we were dealing with future, as well as present, technology, and because we were envisioning, rather than actually designing a system, we did not take the time to weave the technologies together into a coherent design. Rather, we simply defined the categories of technology that might be necessary and the kinds of technology that might appear in each category.

A. Input Technology

The goal of our input technology is to allow any form of information (e.g., text, image, sound) to be sent into the System by any type of person. In other words, regardless of a person’s technical literacy, he or she can still submit information to the System. In some cases, the System will employ so called “adaptive” technologies to assist people with no or limited sensory inputs to still interact with the System. In other cases, we expect intermediaries to work on behalf of non-computer literate participants to help them digitize print or spoken inputs. Intermediaries might use computer-based audio and text translators or technologies that help them understand the wider cultural context of their clients in order to help them interact successfully with the System.
Testing the Access to Justice Technology Bill of Rights

B. Output Technology

As with the input technology, our goal for output is that it be accessible to any type of person. Thus we envision that the System uses the most appropriate means of communicating. For example, it sends surface mail or e-mail, makes phone calls, outputs to computer screens, or prints materials. In the case of interactions with illiterate participants, the System provides materials that a human intermediary would pass on verbally.

C. Storage Technology

How information is stored in the System will be critical to assuring that it can be input in all its forms, output in all its forms, and made available to the right people at the right time. To appropriately store information, the System uses the most advanced databases, structured content, and access technologies. Database technologies allow massive amounts of information to be efficiently stored and retrieved. Structured content technologies, such as XML, allow information, and the rules of its use, to be precisely defined. Finally, access technologies (e.g., indexes, taxonomies, cross references, thesauri, subject headings, and vocabularies) allow precise and thorough cataloging of information so that it can be found when needed.

D. Processing Technologies

An essential feature of the System is process management. Given a set of rules of legal conduct, and a set of participants in a legal issue, the System orchestrates the actions of the players to be consistent with rules of conduct and to move toward a defined type of resolution. The System manages who will accomplish what tasks and at what time.

To accomplish this complex task, the System must first be able to read, interpret, and apply the rules of legal conduct. Just as importantly, however, the System must be able to teach each participant about the process he or she is to follow. The System must empower participants with information not only about the choices they face, but with the probable outcomes of such choices. In other words, the System is capable of recommending a hypothetical solution and then illustrating the probable consequences based on the choices that participants make.

Workflow tools more advanced than any currently available would be employed to manage this sort of process. A language of legal process in
which the logic of automated options for each type of legal issue would be developed. Finally, the feedback loop from human-generated options to the automated logic would be established.

Critical to the success of the System’s scenario building software will be our ability to monitor the past choices that participants have made, and their outcomes, to provide reasonable projections of the future. So, for example, if a particular choice has led in ninety-eight percent of the past cases to a particular outcome, that fact ought to be available to people faced with that choice today. A wide variety of pattern-seeking and statistical analysis technologies would create this capacity.

E. Connections

The System is connected to a large array of other systems to which it provides and receives information (social services, immigration, financial, educational, and many other information sources). The System is able to easily interact with these other systems while still maintaining confidential financial information and providing access to public financial information.

F. Accountability Technologies

Within the rules governing privacy, the general public has access to the System in order to view information regarding the cases within it. The System is open to the public for review yet maintains individual’s anonymity and confidentiality. In addition, the participants and the public submit evaluations concerning how they feel the System handles cases. The participants respond to System generated evaluative questions to provide constant feedback to allow ongoing improvements to the process.

G. Privacy Safeguards

Safeguarding the privacy of information is a primary concern of the System. The System uses every available means to assure that only those authorized by the legal process can see a piece of information. There are two separate classes of technology involved in the creation of safeguards. First, there are the technologies that secure information. Encryption and authentication technologies are chief among the ways that information is made inaccessible. Second, there are the technologies that match a person to the information that he or she may access. These
technologies are built chiefly on access lists and rules processing software. The rules that are processed are, in this case, the laws that govern privacy and access.

In addition to safeguarding access, the System safeguards data from catastrophic loss. Thus, the technologies behind backups, self-correction, and fail-safe are also quite important to the System.

VI. SYSTEM DESIGN INSIGHTS, OPPORTUNITIES, AND BARRIERS

A. General Observations

1. The Changing Role of the Judiciary

Prior to the 1970s the primary job of the courts in the United States was to be finders of fact. The ideal was to be impartial and nearly devoid of intimate contact with and knowledge of litigants and their circumstances. Since that time, society and its institutions have had to deal with a drastically different definition of family. As the traditional family has become less common, institutions, particularly schools, have changed and expanded their roles. Rather than simply being educators, teachers, counselors, and coaches, schools find themselves addressing more than the educational needs of children. The curriculum within kindergarten through twelfth grade has significantly changed as a result.

When schools are unable to deal with the myriad of problems our children face, the court system now steps in to partner with the schools in a therapeutic response to the problems of families and their children. Today we have drug courts, domestic violence protection calendars, juvenile diversions, mandatory coping with divorce classes, and other programs to deal with the breakdown of families that go well beyond the detached fact-finder and reactive neutral decision maker.

The System we envisioned assumes that our justice system would embrace both the older fact-finding mission as well as the newer therapeutic mission of the courts. In addition, we added a new mission—to allow people to come to their own resolution where possible by providing as complete and understandable a system as possible for defining and exercising legal options.
2. The Changing Role of the Legal System

The current and coming generations of information technology enable our System to collect and distribute high quality information. Participants need to define and resolve their differences by themselves. Through technology, the System empowers and enables participants to make well-informed and deliberate decisions. The increased quality and quantity of information available to participants enhances their decisions and thus enables both a greater quality and equality of justice. The System is expected to move the participants through the process quickly, but without haste, hastening the transition in people’s lives from the uncertainty their legal issue creates to the certainty that a resolution is possible and that its outline is clear.

In summary, the System is different in a number of critical ways:

a. Self-Adjusting Interface System

The System’s interfaces are self-adjusting. This concept means that the System does not necessarily accept the participant’s self-definition of his or her role, capacities, and needs, but also measures the way the participant is interacting with the System and adjusts accordingly. This concept makes it possible for the System to be much more “appropriate” for each participant.

b. Multiple Personality System

To further enhance the effectiveness and performance of the System, a multiple personality system was developed. This concept means that the System actually performs a variety of tasks while playing a variety of roles. The System is a neutral information provider. When appropriate it becomes a supportive and zealous advocate during the imposed resolution phase. It may also be a neutral predictor and mediator. Finally, it is a neutral decision maker or adjudicator.

c. Flexible Dynamic Monitoring of Progress

As individuals utilize the System, it constantly monitors the way participants are moving toward resolution. Depending upon the exercise of their options, it changes the way it treats the participants. It also monitors the speed at which the case is moving through the System. If the problem is not being resolved expeditiously, the System will become more aggressive. After the participants reach impasse, the System will
Testing the Access to Justice Technology Bill of Rights

then move out of stasis and, if necessary, impose a resolution in a timely manner. The System’s decision to determine that impasse has been reached will be, in part, statistically driven. The algorithm will not only include the absence of current movement but whether there is a longer-term pattern of interaction and movement. Flexible dynamic monitoring of progress will result in many cases reaching points of appropriate intervention or resolution more quickly.

Such an algorithm-driven system creates a risk of injustice to participants if the individual is forced to particular behavior based upon prior behavior of others or the lack of behavior by the opposing participant. Declaring impasse cannot only be tech driven. Again, human judgment must enter into the decision to declare impasse.

d. Ongoing Analysis of Patterns and Outcomes To Facilitate Resolution

We envisioned a System that automatically performs an ongoing analysis of patterns and outcomes in all cases. This concept is expected to facilitate resolution by optimizing the types of options presented to the participants along with the statistically generated probable outcomes associated with each option. In other words, ongoing data from each interaction feeds into a database that is then used by the System not just to observe and predict outcomes of various options, but to review and modify the options suggested by the System.

The nature of the System according to this concept is purely statistical. The opportunity for improved outcomes as a result of the ongoing analysis of patterns is real. However, human creativity must still be permitted to avoid outcomes, which may not be optimal for the participants.

B. Dangers of the Process and Insights

In Thomas More’s Utopia there are no lawyers because the laws are simple enough for the common person to understand and apply. Our System follows a similar route. It attempts to create an interface to the legal system that any citizen can understand and use. On the other hand, we do not believe that any interface can replace the position of people in the System.

To ensure equality of access to the System, persons are necessary to provide adequate notice. Intermediaries are essential to assist individuals in accessing the System. Mediators must exercise discretion and
creativity in guiding participants to a resolution of their problem. Adjudicators must exercise discretion to determine what decisions may be resolved electronically, through automation, and which decisions will necessarily involve an adversarial hearing. Advocates are necessary to assist individuals using the System. Finally, the participants themselves must play a role in the interpretation of the information and options the System provides.

In short, technology can be a tool that not only makes the System function more efficiently, but also enhances the value of outcomes and, in turn, people’s lives. However technology will not replace people in the System.

Some of the risks we identified as inherent in our System were:

*Funding.* The cost of funding such a system must be resolved. Is it a vision that is simply too expensive? Is it too complex?

*Participants may not understand their options.* What if participants are paralyzed by the number and complexity of options that the System offers? What if, instead of seeing the System as empowering, they see it as an encumbrance because it requires them to do too much of their own learning and analysis? We assume that in the worst case, someone can seek the assistance of a legal professional and take on no more of the burden of understanding than they do today. Still, today it is assumed that people will need legal professionals and they can be found at a low cost. If the System allows persons with enough motivation to find their own way to the resolution of an issue, will low or no cost legal aid disappear?

*The right interface.* We presume that participants will willingly use a system that has interfaces that meet the needs of all. However, a significant amount of public education and marketing will still be necessary to build trust in those who should be accessing the System. Is it possible and economically feasible to build such a universal interface?

*Role confusion.* There is the risk that participants will not understand the different roles that the System takes on (advocate, neutral information provider, and so on). Regardless of the way the System presents itself, they may see it as monolithic and “not on their side.” In addition, the roles themselves will not be clearly defined initially or may blur over time in a participant’s mind.
Testing the Access to Justice Technology Bill of Rights

Trust. Building public trust and confidence in the System will be especially difficult considering the concentration of power in the System. Information is power and having almost all information creates almost absolute power. Coupled with the fact that the System will have the power to impose resolutions as well as enforce compliance, will participants be discouraged from using the System? The amount of personal and financial information the System maintains could compromise public trust. Those who use the System will have fears of invasion of privacy and identity theft.

Intermediaries. The quality of individuals who serve as intermediaries between the System and participants can affect the use and adherence of the System. The behavior and training of these individuals will have to be monitored to prevent incompetence or abuse of power.

Breakdowns. Despite the System’s power, it also has characteristics of vulnerability. If power to the System shuts down, the System shuts down. Terrorists and hackers will be constant dangers to the operation of the System. Flaws and bugs in the software of the System can also cause problems. The System must also have the capability of allowing participants to correct the information it contains in a fair and efficient manner. The use of technology in our idealized System can provide efficiencies and increase the quality of decision-making by its users. However, safeguards must be built into the System to avoid abuse.

Automation. Our System depends on the existence of advanced algorithms for continually getting better at option generation and presentation. It depends as well on algorithms for pattern matching in a database of past issues. These algorithms do not now exist. Are they possible? Can they be created in a robust enough way to be the basis of an entire legal system?

Profiling. People who do not believe that they really match the pattern the System says they do may characterize pattern matching technology as “profiling.” Does pattern matching infringe on individuals’ right to be treated as individuals, rather than as demographic stereotypes? Is the accumulation of information into databases inherently unfair because it results in the transfer of information from the many to the few?
Convergence or self-fulfilling prophecies? Our System attempts to pull participants toward resolution by carefully presenting options, feedback, and scenario-building tools. Will the use of this sort of technology, which provides participants information about probable outcomes based upon their possible choices, lead to self-fulfilling prophecies?

VII. CONCLUSION

To cast the widest conceptual net, our committee intentionally began this study with a small set of goals and no constraints on the technology or processes that might be needed to reach those goals. Our intent was not to design a replacement legal system for the United States, but rather to explore what fair and equal justice might actually mean if all available and imaginable technology were put into practice.

As a result described above, the System we envisioned has many practical problems. An Access to Justice Technology Bill of Rights that takes into account these concerns is but a first step towards creating an ideal civil justice system accessible to all.

We believe that our culture's newfound ability to envision and then build massive information systems is a model for our justice community. If justice is served by open and equal access to the facts and opinions surrounding a legal issue, then an information perspective can help. If such an information system can allow people to solve the majority of their own legal issues and require adjudications in only a small number of cases, then there is hope that when adjudication is needed, it can be given the time and energy it deserves.

The design of this "ideal" system has demonstrated both the potential benefits and dangers of technology. It is the task of all to optimize the potential benefits and minimize the risks. The Access to Justice Technology Bill of Rights can be a crucial tool in doing so.
THE COMMON LAW PROCESS: A NEW LOOK AT AN ANCIENT VALUE DELIVERY SYSTEM

Dennis J. Sweeney*

The life of the law has not been logic: it has been experience. The felt necessities of the time, the prevalent moral and political theories, intuitions of public policy, avowed or unconscious, even the prejudices which judges share with [others], have had a good deal more to do than the syllogism in determining the rules by which [all] should be governed.1

I. INTRODUCTION

This Article is prompted by the Access to Justice Technology Bill of Rights (ATJ-TBoR) project.2 Part of the charge to the ATJ-TBoR Committee is to

[d]evelop and implement an Access to Justice Technology Bill of Rights (“ATJ-TBoR”) premised on relevant principles contained in the United States and Washington State Constitutions, the mission and underlying principles and declarations generating the creation and operation of the Access to Justice Board, the principles contained in the Hallmarks of an Effective Statewide Civil Legal Services Delivery System adopted by the Access to Justice Board in 1995, and subsequent and effectuating documents and declarations.3

One important aspect of this project is the application of the principles generated by this effort to the societal disputes they are intended to influence. Put another way, once we develop these principles, what do we do with them? One goal the ATJ-TBoR project must address is the practical problem of moving the lofty notions

---

* Judge, Division III of the Washington State Court of Appeals. I thank Judge Donald Horowitz, the Chairperson of the ATJ-TBoR Committee, for allowing me to be a small part of this interesting and important project.


generated from the thinking and formulation stage to the practical arena of everyday conflict resolution.

I am a common law judge.4 I serve on a common law court in Washington state. So I spend my days in the realm of what is: real people with real controversies come to the court for the answers to real questions.5 Common law appellate courts answer the questions presented through the vehicle of a written opinion, resolve the dispute, and, incidentally in the process, make law.6

And it is in this sense that common law courts become a delivery system—a delivery system for society’s values. These values in some fashion embody the customs of the community, the society, the courts, and the legal community. As with any set of broad principles, such as those found in constitutions or other basic laws, the ATJ-TBoR principles do not come alive and have no meaning until they are applied to a societal dispute—an actual case or controversy. No one can anticipate all the questions, all the nuances that real people with real conflicts will bring to their application. Until then, they remain largely an abstraction.

It is, therefore, somewhat incongruous for a common law judge like me to participate in a project to develop a set of principles to influence future, and as yet, hypothetical disputes. That conflict came crashing in on me as I struggled to collect my thoughts for this Article. Here, I participate in a project to construct and then promote a set of principles, all without the benefit of a case that has been prepared and argued. The courts of this state have had and will have almost nothing to say about the development of these principles—not, at least, in any traditional common law sense. And so it may be appropriate for me to express some ambivalence about the ATJ-TBoR project, or at least my involvement in it. That said, I want to address technology, societal values expressed as

4. See R.C. VAN CAENEGEM, AN HISTORICAL INTRODUCTION TO PRIVATE LAW 3 (D.E.L. Johnston trans., 1992) ("[T]he common-law system differs fundamentally from the continental system."). The common law is generated by the courts and traces its origins to the royal courts of England beginning in the twelfth century. Id.
5. See R.C. VAN CAENEGEM, THE BIRTH OF THE ENGLISH COMMON LAW 88 (2d ed. 1988) (noting that the common law depends on precedent and is empirical). And, I would add, common law is essentially inductive while continental law tends to be more theoretical and deductive, and based more on abstract principles.
6. EVA H. HANKS ET AL., ELEMENTS OF LAW 4 (1994) ("Unlike statutes and constitutions, the common law rests on no authoritative text external to the judiciary. The law is knowable only by reading past ‘cases’; it is not to be found anywhere other than in those very cases (and in nonauthoritative summaries of them).").
“rights,” and common law courts and processes as a vehicle for transforming values into rights.

The ATJ-TBoR is the product of a group of smart people periodically gathering around a table for a few years and developing what we think are good ideas—based maybe on some custom, legal lore, constitutional values, or case law. Indeed, the history of both common and civil law is a history rich in its diverse sources. But, whatever the source of these principles, the implicit notion is that they are good ideas and should, therefore, be law. In fact, the implicit argument in this project is that the ideas embodied by these principles are so important that they should assume the status of guiding principles for integrating technology into the many facets of modern conflict resolution. Surprisingly, no one questions this approach to the development of broad legal principles.

Have common law courts subtly and incrementally put themselves out of the substantial and traditional business of law-making or, at least, put themselves out of business as we once knew it? More personally, do I belong here? Or am I helping to betray the common law tradition I preach and practice and which has served the citizens of Washington since statehood and before?

The short answer is: I think not.

Each of the relevant principles, if they are to have any practical application at all, must someday be applied by a court to an actual case, to an actual controversy. I suggest in this Article that regardless of the form these principles ultimately assume—court rule, legislation, or uniform code—it will be through common law courts and common law processes that they ultimately become part of the fabric of our legal culture. Only when a trial judge in King, Clark, or Benton County, Washington, treats them as law and applies them as law in the process of deciding a case will they become “law.”

The essence of the common law tradition is a process for resolving societal disputes—a process that generates law as a byproduct. The process does not set out to make law. Its goal is to resolve a particular societal dispute in some principled and structured way. But, in doing so, every decision becomes part of the fabric of “the law.” Every decision

7. See VAN CAENEGEM, supra note 4, at 2 (“Historically, the major elements of the law belong to a common European inheritance: ancient and medieval Roman law, canon law, old Germanic law, feudal law, medieval municipal law, the natural law of early modern times.”).

by a common law court innervates and gives meaning to what otherwise would be just a lifeless paragraph in a dusty book.9

II. TECHNOLOGY AND COMMON LAW COURTS

The common law develops slowly and uncertainly. If no case is brought, no rule is developed. And the common law is conservative. It always looks to what has been done before.10 But technology is neither slow nor conservative. The potential for technology to influence the courts, and the need for the law to accommodate this change, are not unanticipated:

As we look ahead we may reasonably expect great changes in the law, both statutory and decisional, in a number of areas. Illustrative changes include a continued effort to improve the quality of life. . . . Related problems involve . . . increasing demand by many members of the public for greater participation in the benefits of our improving technology.11

But who could have predicted the speed with which technology would influence the entire spectrum of the law12—legal policy,13 trial tactics

10. See Windust, 52 Wash. 2d at 36, 323 P.2d at 243 (noting that “common law is comprised of that body of court decisions in the nonstatutory field to which the doctrine of stare decisis applies”). But see State v. McCollum, 17 Wash. 2d 85, 107, 136 P.2d 165, 174 (1943) (noting that where personal liberty is at stake, stare decisis will not save a prior holding “so patently erroneous that it is hardly necessary to discuss it”).
12. See, e.g., Lotus Dev. Corp. v. Paperback Software Int’l, 740 F. Supp. 37 (D. Mass. 1990) (concluding that a menu command structure of a computer program, including the choice of command terms, the structure and order of those terms, their presentation on the screen, and the long prompts could be protected by copyright law); State v. Townsend, 105 Wash. App. 622, 630, 20 P.3d 1027, 1032 (2001) (stating that use of ICQ computer communications program results in consent to recording by the recipient as a matter of law because of the characteristics of the software), aff’d, 147 Wash. 2d 666, 57 P.3d 255 (2002).
13. See, e.g., Ruth Gavison, Privacy and the Limits of Law, in COMPUTERS, ETHICS & SOCIAL VALUES 332 (Deborah G. Johnson & Helen Nissenbaum eds., 1995); see also Dean Colby, Conceptualizing the “Digital Divide”: Closing the “Gap” by Creating a Postmodern Network That Distributes the Productive Power of Speech, 6 COMM. L. & POL’Y 123, 163 (2001). Colby argues that the mere passive access to modern communication technology is not enough: “[t]he linchpin for any solution to the DD [digital divide] must include a judicial review of the First Amendment that will yield case law that allows the government to protect speech by requiring private enterprise to provide access to the means of speech.” Id.
and legal procedure,\textsuperscript{14} and the courts themselves\textsuperscript{15}—all of which have the potential to influence access to justice? The analogy of law as the tortoise and technology as the hare is an apt one.\textsuperscript{16}

So the need to import values into the judicial process for application to everyday conflicts generated or complicated by technology was apparent. And part of the problem is access. Should the poor, the disadvantaged, be assured access to the courts on an equal footing with the well-to-do members of our society? If they should, then in what way should they participate? And if the inexorable advance of technology can serve either to impede or advance access to the courts, then what can the courts do to ensure that technology does not impede but does in fact advance access?

\textsuperscript{14} See, e.g., \textit{Fed. R. Civ. P. 5}, 43(a); Richard Zorza, \textit{Re-Conceptualizing the Relationship Between Legal Ethics and Technological Innovation in Legal Practice: From Threat to Opportunity}, 67 \textit{Fordham L. Rev.} 2659, 2659–63 (1999) (“For the first time since the invention of the typewriter and the telephone, technology has again begun to exert a significant influence upon the practice of law. New communications technologies hold the promise of increased access to legal services for the public at large, and particularly for the poor. The tide of innovation has, however, triggered disquiet among experts in legal ethics, particularly in response to the practice innovations that incorporate these technologies.”); \textit{see also} Marc A. Ellenbrogen, \textit{Note, Lights, Camera, Action: Computer-Animated Evidence Gets Its Day in Court}, 34 \textit{B.C. L. Rev.} 1087, 1087–90 (1993).


Unequal access to technology may exacerbate the potential mismatch between an affluent client’s technology-savvy counsel and the poor or pro se litigant’s pedestrian procedures. Lederer, \textit{supra}, at 832. But courtroom technology is expanding access for the disabled. For example, real-time transcription and closed captioning allow hearing-impaired people to be jurors, trial counsel, and even judges. \textit{Id.} at 834.

Society changes, culture changes, and as this Symposium and, indeed, the whole ATJ-TBoR project assumes, technology will drive or at least influence that change. But if access to justice is to be a reality, regardless of the inexorable march of technology, then where will the guiding principles come from? For those of us participating in the ATJ-TBoR project, the answers are to be found in the principles generated by this project. The introduction of this basic set of tools is necessary if the courts are to cope with the impact of technology and its ramifications on the justice system. It is then the goal of the ATJ-TBoR Committee to provide that framework.

A. Established Rights and New Technology

Describing the ATJ-TBoR as the development of a list of rights is a bit misleading. Many of the rights in jeopardy, or at least at issue, are well developed and, indeed, already drive courts’ decisions.

Both Washington and United States courts have long and well-established precedents supporting fundamental rights, such as the right to privacy, the right to travel, and the right to a fair trial. The

17. “It is not technology, as such, which affects society for good or bad, but its uses, which are . . . shaped by the values of society and by the historical context in which the technology is used. . . . We must remember that we are not trapped helplessly in front of an unstoppable technological steamroller. Our control is in how we use our knowledge that we will be required to live with the results of our decisions on the use of this new technology.” Peninsula Counseling Ctr. v. Rahm, 105 Wash. 2d 929, 948, 719 P.2d 926, 936 (1986) (quoting Fred W. Weingarten, Privacy: A Terminal Idea, 10 HUM. RTS. Q. 18, 56 (1982)).

18. And, indeed, at least one recently adopted court rule accommodating the use of briefs on CD-ROM appears to be based squarely on the right proposed by Principle Number 1 of the ATJ-TBoR “Requirement of Access to Justice.” Proposed Principle Number 1 would require that “[i]ntroduction of technology or changes in the use of technology must not reduce access or participation and, whenever possible, shall advance such access and participation.” See ATJ-TBoR, supra note 2. Washington Rules of Appellate Procedure permit the submission of a brief on CD-ROM. WASH. R. APP. P. 10.9, available at http://www.courts.wa.gov/court_rules/?fa=court_rules.display&group=app&set=RAP&ruleid=app10.09. But the rule does not require such a format. Id. 10.9(d). Nor does the rule allow the imposition of the costs incurred in the preparation of those briefs. Id. 10.9(f).

19. Griswold v. Connecticut, 381 U.S. 479, 482–86 (1965) (recognizing a right to privacy and also noting that although the First Amendment does not specifically mention the freedom of people to associate, to choose their children’s schools, or choose a particular subject of study, these rights have been construed to fall within the First Amendment); State v. Jackson, 150 Wash. 2d 251, 259–60, 76 P.3d 217, 222 (2003) (recognizing a broader privacy interest under article I, section 7 of the Washington State Constitution than under the Fourth Amendment to the United States Constitution).

challenge to the ATJ-TBoR Committee was to craft a set of principles that will at least safeguard those rights as technology advances. So this exercise of crafting a technology bill of rights entails protecting or, in some cases, revisiting well-established rights such as privacy.

The fundamental right to privacy is well established independently of any consideration of technology. But sophisticated technologies have affected the problem of privacy and, necessarily, the way privacy issues are perceived by the courts. The ATJ-TBoR principles assure that existing and emerging technologies do not undermine this and other already well-established rights.

The courts have already had to grapple with the influence of technology on privacy in a variety of contexts. By statute in Washington, it is unlawful to record any

\[\text{private communication transmitted by telephone, telegraph, radio, or other device between two or more individuals...[using] any device electronic or otherwise designed to record and/or transmit said communication regardless how such device is powered or actuated, without first obtaining the consent of all the participants in the communication.}\]

In three cases consolidated as *State v. McKinney*, each defendant was arrested after police officers accessed information contained in their Department of Licensing (DOL) driver’s license records. Each moved


\[\text{23. See, e.g., Jackson, 150 Wash. 2d at 264, 76 P.3d at 224 (holding that police installation of Global Positioning System (GPS) tracking device on murder suspect’s vehicle infringes upon Washington state constitutional right to privacy and therefore requires a search warrant); State v. McKinney, 148 Wash. 2d 20, 32, 60 P.3d 46, 51 (2002) (evaluating right to privacy in the context of Department of Licensing records); State v. Young, 123 Wash. 2d 173, 186, 867 P.2d 593, 599 (1994) (holding that warrantless thermal imaging to detect heat sources within home is a violation of right to privacy); State v. Gunwall, 106 Wash. 2d 54, 63, 720 P.2d 808, 813 (1986) (holding that police was not permitted to learn who a person was contacting by tracking her phone calls); State v. Myrick, 102 Wash. 2d 506, 511, 688 P.2d 151, 154 (1984) (holding that warrantless aerial surveillance violates state privacy clause).}\]

\[\text{24. WASH. REV. CODE § 9.73.030(1)(a) (2002).}\]

\[\text{25. 148 Wash. 2d 20, 60 P.3d 46 (2002).}\]
for dismissal or suppression on the ground that the search of DOL databases prior to arrest violated the privacy provisions of the state constitution. Their motions were denied, and the defendants were convicted. The court held that access to a computer database to confirm license plate numbers is not an invasion of the vehicle owner’s right to privacy:

Based on the historical treatment of driver’s license records, the fact that these records reveal little about a person’s associations, financial dealings, or movements, and the purpose for which the State compiles and maintains these records, we hold that there is no protected privacy interest in the information contained in a DOL driver’s record under article I, section 7 of our state constitution.26

Likewise, the release of the names of public employees, without more, has been held not to be a violation of privacy, even if that information could be linked to other computer-generated information about the employee.27 An essential argument in both *McKinney* and *Tacoma Public Library v. Woessner*,28 was that technology—the computer—with its capacity to collate information easily and quickly should change the rules for disclosure of information (licensing information in *McKinney*, and personal information in *Woessner*) given the statutory and constitutional right to privacy in Washington. The courts rejected arguments for nondisclosure simply because computers changed the ease and accuracy with which this information could be associated or accessed.29

But evidence obtained in violation of Washington’s privacy statute30 is inadmissible in a criminal case.31 The Washington State Supreme

26. *Id.* at 32, 60 P.3d at 52.

27. King County v. Sheehan, 114 Wash. App. 325, 346, 57 P.3d 307, 317–18 (2002) (“It is a fact of modern life in this age of technology that names can be used to obtain other personal information from various sources, but we conclude that this is not sufficient to prevent disclosure of the names of police officers under the act.”); Tacoma Pub. Library v. Woessner, 90 Wash. App. 205, 218, 951 P.2d 357, 363 (1998) (involving release of names of library employees).


31. *Id.*
Court concluded that under Washington’s privacy act, e-mail messages to a fictitious minor to arrange for an illegal sexual liaison were private and hence inadmissible because the communications in real time by way of the ICQ instant messenger program were intended to be private and were recorded on a device (the sting officer’s computer).32

Similarly, in State v. Nordlund,33 the appellate court noted with approval the trial court’s characterization of the “personal computer as ‘the modern day repository of a man’s records, reflections, and conversations’” and thereby conferred Fourth Amendment protection to the exercise of a First Amendment right.34

Suits or defenses based on privacy are nothing new—and, indeed, may be on the increase as a byproduct of advancing technology. The object here is to provide the courts with principles that assure that technology, and the advance of technology, do not erode fundamental rights, like privacy.

B. Implied Rights

I will make another point on the topic of rights and common law courts. I have selected rights that are, at least constitutionally, implied. I will explore this more fully in my review of Washington case law supporting access to justice. But my point here is that it was necessary for the courts to imply these rights in order to give effect to other express and accepted rights.

To take a few other examples: The right to vote—how do I participate in democratic government if my right to vote is restricted? The right to work—how do I work (particularly in this modern economy) if I cannot travel? The right of access—how do I assert any right or vindicate any right for others if I am denied access to the courts because of poverty? And, finally for purposes of the ATJ-TBoR, how do I participate in any dispute resolution process when I neither have access to nor understand the dizzying array of technology that is fast becoming a part of the system?

34. Id. at 181–82, 53 P.3d at 525 (quoting trial court record).
III. COMMON LAW

I turn now to a few observations about the operation of a modern common law system, or at least, the operation of the common law system in Washington state.

A. Common Law Reasoning

The courts of this state exercise their authority through legal reasoning. They do not provide a technique for supplying the right answers.35 Indeed, in this pluralistic society, the “right answer” is frequently the subject of some debate. So “ultimate results are often less important than how they are arrived at.”36 In fact, “[t]he most important ‘reception’ from England was perhaps not any particular body of doctrine but a way of thinking about law which made it easier for our great jurists (such as Story) to create an indigenous American jurisprudence.”37 What the courts do, instead, is choose between competing arguments and then, through the vehicle of a written opinion, justify that choice.

And the court’s analysis can always be distilled into a logical syllogism.38 The major premise is the statement of the law—the rule: running a red light is negligent; obtaining money by lying is fraud. The minor premises are the particular facts of a case: the defendant ran the red light; the defendant lied to obtain money. And, finally, the conclusion: the defendant ran the red light and was, therefore, negligent. Or the defendant lied to obtain money and was, therefore, guilty of fraud.39

In modern common law analysis, the major premise—the rule—frequently originates outside of the judicial system. The ATJ-TBoR was not generated by our common law court system. And that may well be in keeping with the modern common law approach to the development of law. So then, where do these rules come from?

37. HANKS, supra note 6, at 7.
38. Perhaps the most familiar of all syllogisms: All men are mortal; Socrates is a man; Socrates is therefore mortal.
Common Law Process

**B. The Major Premise (the Rule) and Its Sources**

“What is it that I do when I decide a case? To what sources of information do I appeal for guidance? In what proportions do I permit them to contribute to the result? In what proportions ought they to contribute?”

Any common law resolution of a case begins with the identification of some principle—some rule of law. And a variety of sources contribute to those rules. It is not usually a judge-made rule that adds the major premise to the syllogistic proposition that ultimately resolves the case. Sources include state and federal constitutions, state and federal legislation, learned treatises, law reviews, and restatements of the law. But all have a common attribute. They have been identified by some court as worthy of application to a societal dispute. In traditional common law, the rule with which the court started reflected the custom, or at least the court’s understanding of the custom of the realm. Indeed, the common law system has been described as the “custom of following custom.” And, at least during one period of the long common law tradition, the rules articulated by the judges were, if not God’s law, at least a reflection of the laws of nature.

We have, of course, abandoned any notion of judges speaking eternal truths in favor of a more rational, utilitarian approach to values, at least those that we are willing to apply to societal disputes. But, that said, the grounding of common law values in natural law endowed the system with the aura of doing the “right thing”; some sense of the absolute; something beyond the here and now and the “just us.” This notion seems to have been accepted when our state constitution was adopted. And I think it finds expression in Washington’s access to justice cases.

41. See ALDISERT, supra note 39, at 28.
42. CARTER, supra note 35, at 121–22.
43. 1 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 42 (1983); CARTER, supra note 35, at 123–24; RICHARD A. POSNER, THE PROBLEMS OF JURISPRUDENCE 12 (1990); M. Stuart Madden, The Vital Common Law: Its Role in a Statutory Age, 18 U. ARK. LITTLE ROCK L.J. 555, 572 (1996) (“The origins of the common law can be traced at least from Aristotle and Cicero through the Book of Exodus. It is generally supposed that much of the animating basis for early common law derived from an innate, elemental, and sometimes theocratic concept of justice often termed ‘natural law.’”).
That aside, the major premise, the black letter rule, of a court’s syllogistic exercise is frequently supplied by sources outside of the court system.45 Of course, in a democratic and remarkably populist state (by constitutional declaration of rights, all political power resides in the people),46 much of our public policy is outside the province of the courts. The usual source is the legislature. We now have extensive codes for everything from commerce,47 to crime48 and sentencing,49 to probate,50 and marital dissolution.51 Uniform laws fill almost every nook and cranny of areas previously reserved to the discretion of judges, the courts, and case precedent.

But, regardless of the source of the general principle or major premise, the black letter rule reflects the custom of the society. Or, at least, it represents the ideal of what that custom should be. A common law system is said to identify and apply the “customs” of the community.52 The sources used to identify those customs have changed over the centuries. But the courts still work to identify principles that in some sense continue to represent the customs of a community.

So one obvious and important source of rules for judges in this state is the legislature.53 But it is not the only source. In our modern common law system there are many other sources of law. And each claims its own authority to speak for the community. Each claims its own right to say that this principle, this statement, this rule, is the custom in the community or is generally acknowledged to be a good idea and should be the custom in this community.

Some rules are articulated by the Washington State Supreme Court acting in its rulemaking capacity.54 Some are the product of a persuasive

45. CARTER, supra note 35, at 15.
46. WASH. CONST. art. I, § 1.
49. Sentencing Reform Act of 1981, WASH. REV. CODE ch. 9.94A.
50. Probate and Trust Law, WASH. REV. CODE tit. 11.
52. LAWRENCE M. FRIEDMAN, A HISTORY OF AMERICAN LAW 21 (2d ed. 1985).
53. See State v. Laitinen, 77 Wash. 2d 130, 133, 459 P.2d 789, 791 (1969) (stating that courts do not consider whether a statute plain on its face is unwise or ineffectual, only whether it is within the legislature’s constitutional powers).
54. WASH. REV. CODE § 2.28.150 (“When jurisdiction is, by the Constitution of this state, or by statute, conferred on a court or judicial officer all the means to carry it into effect are also given; and in the exercise of the jurisdiction, if the course of proceeding is not specifically pointed out by
law review or other learned treatises, in which the authors collect, synthesize, and analyze what others have said about a topic. 55 Some rules are the product of restatement editors as they collect and organize what courts have said about an issue. 56 And traditional common law remains by legislative mandate a source of rules. 57 This is particularly so with the contribution of the American Law Institute, which identifies and distills common law rules and systematizes them in the process. 58 Even a cursory review of Washington cases shows that the sources of the principles that control the outcome of societal disputes are many and varied. Finally and specifically, the principles represented in the ATJ-TBoR are founded upon Washington’s tradition of universal access to justice.59 These are principles identified by a broad spectrum of lawyers, judges, scholars, and citizens concerned with the poor, the courts, and technology.60

The wells from which common law courts have drawn the customs of the realm have changed over the years. But identification of these rules, principles, and statements continues to represent an acknowledgement by the courts that a rule, principle, or statement in some way represents the demands of an ordered society. The law generated by appellate judicial opinions should reflect the social, political, and economic influences and ideas of our age.

55. See Del Rosario v. Del Rosario, 116 Wash. App. 886, 893, 68 P.3d 1130, 1133 (2003) ("Bennett has been criticized. The Corbin treatise concludes it is wrong. A law review article makes an extensive analysis in disapproving of the rationale and possible broad holding." (quoting Nevue v. Close, 123 Wash. 2d 253, 256, 867 P.2d 635, 636 (1994))).

56. FRIEDMAN, supra note 52, at 676.

57. WASH. REV. CODE § 4.04.010 ("Extent to which common law prevails. The common law, so far as it is not inconsistent with the Constitution and laws of the United States, or of the state of Washington nor incompatible with the institutions and condition of society in this state, shall be the rule of decision in all the courts of this state.").

58. FRIEDMAN, supra note 52, at 676.

59. See infra Part IV.

60. See supra Part I.
Case law is the wide stream flowing uninterrupted since the twelfth century.\textsuperscript{61} Principles like those set out in constitutions or statutes or those represented in the proposed ATJ-TBoR are, then, the levees, embankments, and even the dams that structure and direct the resolution of that stream of disputes.

The ultimate expression of law by the courts may not always be clear, clean, or even well organized, but that reflects the reality of human conflict. That reality is driven home when one asks: How will the ATJ-TBoR influence a specific divorce action? A criminal prosecution? A complex commercial case? Without the context of a specific case, the answer is unknown and unknowable. It may be educated speculation, but speculation nonetheless.

Through the work of the ATJ-TBoR Committee, an important new source of legal principles has been added to those a Washington judge can draw upon. And, ultimately, the speed at which these principles work their way into the common law of Washington may not depend upon the vehicle used to implement them—legislation, Supreme Court rule, uniform body of law, law review. It may instead turn on the work of judges of this state, who, with the aid of trial and appellate lawyers, identify them in traditional common law parlance as customs of the community.

Let us now move from the major premise, the rule, to other steps in the common law process of deciding cases.

C. Common Law as Process

If we regard the common law as a mechanism for delivering the values of a society at a given time, the common law courts are very much alive. Indeed, they are an indispensable and vibrant dispenser of a society’s values. “The common law is not a body of rules; it is a method. It is the creation of law by the inductive process.”\textsuperscript{62} It is courts that first identify principles from a variety of sources and then apply them by traditional processes based on the exigencies of a given case and, in the process of doing so, make law.

The processes by which the lawsuit and law developed have remained stable, even if the sources for the rules have changed and expanded over

\textsuperscript{61} VAN CAENEGEM, supra note 4, at 3 (noting that the common law developed “from the twelfth century” and is “characterized by historical continuity”).

\textsuperscript{62} Madden, supra note 43, at 559 (quoting RICHARD A. COSGROVE, OUR LADY THE COMMON LAW: AN ANGLO-AMERICAN COMMUNITY 1870–1930, at 33 (1987) (citations omitted)).
Common Law Process

the centuries. Lawyers still bring a case—a question—to the courts. A trial judge in some fashion decides the facts, with or without a jury. Those facts become the minor premises in the syllogism that ultimately dictates the result.

Whether the source is the constitution, a statute, or a prior judicial decision, all rules must be applied before they become law. In at least one sense, then, the rules do not dictate the result. Indeed, in the common law tradition, the function of rules is guidance. Common law processes allow the courts to tailor broad principles and general rules to the demands of the particular case before the court. It is always the general being applied to the particular. “Thus, one possible answer to our initial question (‘what is law?’) is that, at least in ‘common law,’ law is application—application of legal norms by individuals in ordinary interactions.”

Common law processes include, at a minimum, a case in controversy, identification of a general principle, and application of that principle to the specific case in controversy, employing canons of construction and interpretation, standards of review, equity, juries, and a judge exercising discretion at every stage of the process. Even if a case is not strictly speaking one of first impression, the nuances of a specific fact pattern will amplify or explain established precedent. Thus, a new statement of the law results from the resolution of the case.

Even the selection of the rule requires an exercise in judgment, an exercise in discretion by the court. Is the right at issue constitutional or one conferred by statute or court rule? If the rule is a constitutional mandate, is it controlled by our Washington State Constitution or the United States Constitution?

64. See Maurice Rosenberg, Judicial Discretion of the Trial Court Viewed from Above, 22 Syracuse L. Rev. 635, 643 n.19 (1971) (“The justification for discretion is often the need for individualized justice. . . .” (quoting Kenneth Davis, Discretionary Justice: A Preliminary Inquiry 17 (1969))).
65. Hanks, supra note 6, at 4.
the principle? And if so, is it within the purview of the legislative enabling act, and is it constitutional? And even if the principle is supplied by a statute, regulation, or constitution, do the elements meet traditional common law requirements, say, to sustain a conviction for the commission of the crime charged? Classification of a question as arising in equity or at law greatly influences the application of the principle and the ultimate outcome.

A general principle of law is a statement of values. But what does it mean? How is it to be applied? In short, how does the principle find expression and application to the societal disputes it is intended to affect? It is here, as a mechanism by which evolving societal values are transmitted and applied, that the essence of the common law system endures. Values are transmitted by common law courts resolving everyday societal disputes, both civil and criminal.

First and probably foremost, the standard of review determines the degree of deference, if any, to be given to the trial court’s interpretation and application of the rule. Next, the principle may be filtered through one or more canons of construction to determine the manner in which it will be applied. Rules of evidence, whether incorporated into court rules or established at common law, may also influence the outcome of a given case. Outcomes are also affected by conscious and (as Justice

72. See, e.g., In re Parentage of Jannot, 149 Wash. 2d 123, 126, 65 P.3d 664, 666 (2003) (stating that child custody procedural rulings reviewed for abuse of discretion, not de novo, even if based on affidavits); State v. Read, 147 Wash. 2d 238, 243, 53 P.3d 26, 30 (2002) (noting that standard of review for refusal to instruct jury on self-defense depends on reason for refusal).
74. A party who moves unsuccessfully to exclude the opponent’s proposed evidence may offer that same evidence without waiving his claim of error. See State v. Thang, 145 Wash. 2d 630, 647
Common Law Process

Holmes reminds us\textsuperscript{75}) unconscious factors the court brings to bear on the particular dispute before it. Finally, such traditional common law processes as harmless error\textsuperscript{76} and appellate deference to judicial discretion\textsuperscript{77} give common law courts a significant role in bringing general principles to bear on the outcome of a given case.

Besides that, all of this is filtered through minds trained in a specific tradition:

The judge, even when [s]he is free, is still not wholly free. [S]he is not to innovate at pleasure. [S]he is not a knight-errant roaming at will in pursuit of his [or her] own ideal of beauty or of goodness. [S]he is to draw his [or her] inspiration from consecrated principles. [S]he is not to yield to spasmodic sentiment, to vague and unregulated benevolence. [S]he is to exercise a discretion informed by tradition, methodized by analogy, disciplined by system, and subordinated to “the primordial necessity of order in the social life.”\textsuperscript{78}

The common law is, then, a process of courts applying societal norms to people in real life disputes. Judges exercise discretion at every stage of the process.

A court finds the facts of a given controversy. It then decides which of those facts are material and excludes others as immaterial. Only the material facts are then permitted to influence the decision in a given controversy. The court next identifies the legal principle to be applied to that specific set of facts. And deciding that one legal principle controls

\textsuperscript{75.} See supra note 1.


To identify the nature of the trial court’s discretion is not, of course, to identify the nature of our own discretion on appeal. When a trial judge’s function is to decide whether the evidence is sufficient to support a finding, a reviewing court’s function will be the same. When a trial judge’s function is to decide whether the evidence preponderates, a reviewing court’s function may or may not be the same. If the reviewing court’s information is as good or better than the trial court’s, the reviewing court will sometimes be permitted to substitute its own view, without deference to the trial court; but if the reviewing court’s information is not as good as the trial court’s, the reviewing court will limit itself to deciding whether the evidence is sufficient to support what the trial court did.

\textit{Id.} at 104, 971 P.2d at 566 (footnotes omitted).

\textsuperscript{78.} CARDOZO, supra note 40, at 141 (edited for gender neutrality and footnote omitted).
over another is an exercise in judgment by the court. Many dissents are generated by disagreement over the dispositive rule. 79 The court then, of course, applies these principles to the facts. 80

Finally, juries remain an important part of our legal tradition. Their application of a given rule to a specific controversy is nothing less than a reflection of the customs of a society. An intricate, nicely worded jury instruction setting out the rule to be applied may provide a trial lawyer some fodder for appeal. But, once the jury retires to deliberate, it is free to ascribe to the words whatever meaning it chooses. 81

Rumors of the demise of the common law may be premature. 82 Codes, rules, and statutes do not apply themselves. Rather, it is the common law process of making law by arriving at principled decisions in real cases, based not upon a single rule or statute, but upon a whole tradition of looking at and thinking about law and legal problems that both yields a decision and incrementally adds to the body of common law. The common law approach remains the single most effective mechanism for adapting the law incrementally to society’s changing values. Witness the way in which the law has accommodated dramatic changes in our economic structure and practices, 83 or the incorporation of the ongoing scientific advances in DNA evidence. 84 And, of course, the common law approach retains the ability to tailor law to the needs of the individual case.


80. See MEADOR, supra note 9, at 235–36.


82. Madden, supra note 43, at 555 n.2.


Common Law Process

With these reflections on the common law in the twenty-first century here in Washington, let me now turn to Washington’s evolution of the notion of access to justice and examine some of these processes at work.

IV. WASHINGTON’S ACCESS TO JUSTICE

A bedrock foundation for the ATJ-TBoR project and, indeed, for the whole concept of access to justice, is the notion that all citizens of this state have a basic right to access justice. But this principle will not be found in any state or federal constitution, or in any statute, book, or rule. The principle of access to justice will only be found in a short list of Washington State Supreme Court cases beginning with *O’Connor v. Matzdorff* in 1969 and ending with *Doe v. Puget Sound Blood Center* in 1991.

I will focus on two aspects of each case: first, the underlying principle supporting the concept of access to justice in this state; second, and just as importantly for this Article, the source of authority from which the court derives the principle underlying access to justice—the well from which the court drew its inspiration, if not the rule.

A. *O’Connor v. Matzdorff*

Glennie O’Connor tried to file a lawsuit in Yakima Justice Court for replevin and damages of $215.50. State statute required payment of a filing fee. Ms. O’Connor filed a motion and affidavit to proceed *in forma pauperis* instead of paying the fee. The court clerk refused to accept

---


86. *But see Saylors*, 87 Wash. 2d at 739, 557 P.2d at 325 (questioning the existence of any such right, or at least its foundation); *Filan v. Martin*, 38 Wash. App. 91, 97, 684 P.2d 769, 772–73 (1984) (in response to the assertion of a constitutional right to access, the court notes that “[t]he [N]inth [A]mendment to the United States Constitution ensures only those rights deemed fundamental by history and tradition. It does not necessarily give constitutional magnitude to all unenumerated rights. The same analysis applies to the Tenth Amendment.”) (citations omitted).


89. *O’Connor*, 76 Wash. 2d at 590, 458 P.2d at 154–55.
the filing. Ms. O’Connor petitioned the Washington State Supreme Court for a writ of mandamus requiring the clerk to accept and file her papers without payment of a fee.90

The first question the court addressed was whether to assume original jurisdiction and hear the case at all.91 Holding that the question of whether Ms. O’Connor was entitled to proceed with her suit despite her poverty was “fundamental,” the court agreed to decide this question.92

The court based its holding that the right to sue, despite indigence, was fundamental on a series of law review articles and a book.93 A preliminary proposition established was that, at least, the question of poverty as a bar to suit is fundamental.94 The right is then accepted as fundamental in a later case.95

The next question was whether or not Ms. O’Connor was indigent as a matter of law.96 The court concluded that, given the size of her family and that her sole income was a public assistance grant, she was indigent.97 This legal conclusion was based on authority of a Social Security Administration definition of poverty, an editorial in the Seattle Post-Intelligencer, and a law review article.98 The court then expanded and clarified the legal definition of poverty based on an earlier

90. Id. at 590–91, 458 P.2d at 155.
91. Id. at 591–92, 458 P.2d at 155.
92. Id. at 592, 458 P.2d at 155–56.
93. Although only an individual’s right is being asserted in this proceeding, the question to be decided involves very deeply the interests of the public and in particular those of a regrettably large segment of our society. The right of the poor to obtain redress for wrongs, and to defend themselves when sued by the more affluent, is presently of nationwide concern, as is evidenced by the attention given to the subject in legal periodicals. Some notable discussions are to be found in the following: Samore, Legal Services for the Poor, 32 ALBANY L. REV. 509 (Spring 1968); Shriver, Law Reform and the Poor, 17 AM. U. L. REV. 1 (Dec. 1967); Stumpt, Law and Poverty: A Political Perspective, 3 WIS. L. REV. 694 (1968); Silverstein, Waiver of Court Costs and Appointment of Counsel for Poor Persons in Civil Cases, 2 VALPARAISO U. L. REV. 21 (Fall 1967); Barvick, Legal Services and the Rural Poor, 15 KAN. L. REV. 537 (1967). The leading article was written years ago by John MacArthur Maguire, Poverty and Civil Litigation, 36 HARV. L. REV. 361 (Feb. 1923), reviewing the history of the relations between poor people and the courts and lamenting the slowness of the movement toward justice for the indigent. See also a book by J. COMER, FORGING THE FEDERAL INDIGENT CODE (1966).
95. O’Connor, 76 Wash. 2d at 593, 458 P.2d at 156.
96. Id. at 593–94, 458 P.2d at 156–57.
97. Id. The law review article was William Samore, Legal Services for the Poor, 32 ALBANY L. REV. 509 (1968).
Common Law Process

Washington State Supreme Court decision99 and a United States Supreme Court decision:100

We held that the term does not and cannot, in keeping with the concept of equal justice to every man, mean absolute destitution or total insolvency. Rather it connotes a state of impoverishment or lack of resources on the part of the defendant which, when realistically viewed in the light of everyday practicalities, substantially and effectively impairs or prevents his pursuit of his remedy.101

The next question in the analysis was “has the court the inherent power to waive fees prescribed by statute”?102 In dicta, the court observed that if a court had created the rule requiring the filing fee, a court could waive it. But the fee here was required by legislative enactment. The court ultimately held that it has “inherent power” to waive the fee, notwithstanding the statute.103 The source for this remarkable holding is important. The court quoted extensively from an American Law Report, which in turn relied on ancient treatises (Britton: Views on Disseisin; Mirror of Justices); English case law; English statutes (from the reigns of Henry VII and Henry VIII); and miscellaneous state and federal cases.104 The court also concluded that justice courts have the same inherent power because “[t]he justice courts of this state do exercise a portion of the common law jurisdiction.”105

The final question was whether the court “should exercise the power to waive its fees in a given case.”106 And the answer to that question turned on the validity of the argument that the poor, unchecked by prohibitive filing fees, would “inundate the courts with frivolous cases.”107 And, again, the authority for the court’s conclusion that the concern was unwarranted included two law review articles, one of which cited a legal periodical.108

102. Id. at 597, 458 P.2d at 158 (emphasis added).
103. Id. at 600, 458 P.2d at 160.
104. Id. at 598–600, 458 P.2d at 159–60.
105. Id. at 604, 458 P.2d at 162.
106. Id. at 600, 458 P.2d at 160.
107. Id. at 601, 458 P.2d at 160.
108. Id. at 602, 458 P.2d at 161–62.
With *O’Connor*, access to justice became a reality in Washington. In sum, the court decided: (1) the right of an indigent citizen to sue is fundamental and significant; (2) “indigent” does not necessarily mean “utterly destitute” or “totally insolvent”; and (3) courts have inherent authority to waive a filing fee imposed by the legislature.

Another interesting and significant aspect of this case is the court’s approach to identifying and supporting the principles announced. It is classic common law. What is the custom of the community on these questions? The answer, if *O’Connor* is any guide, is that it is what lawyers, commentators, and judges are saying and have said it is and the court’s own sense of the felt necessities of the times.109

B. Iverson v. Marine Bancorporation

In *Iverson v. Marine Bancorporation*,110 Lil Iverson recovered a $1000 judgment following a wrongful eviction action. She did not think it was enough, so she tried to appeal.111 Following some procedural wrangling, the Washington State Supreme Court remanded to the trial court for findings of fact on those questions required by the *O’Connor* decision. A trial court then found (1) Ms. Iverson did not have the money to perfect and prosecute an appeal, and (2) her appeal was prosecuted in good faith and was not frivolous.112

With that factual backdrop, the question framed for the court was whether the Washington State Constitution obligated the court to accept the appeal without cost.113

The court began its analysis with article IV, sections 1 and 30 of the Washington State Constitution.114 These provisions vest all judicial

111. *Id.* at 164, 517 P.2d at 197.
112. *Id.* at 166, 517 P.2d at 198–99.
113. *Id.* at 166–67, 517 P.2d at 199 (“[Amicus] argue that the state constitution sets up both the Court of Appeals and this court to handle a particular mission. That mission, they argue, is for the courts to hear and decide all cases regardless of whether the parties are rich or poor; that they be accessible to all citizens; and that they resolve individual and social conflicts regardless of whether the parties are rich or poor. We agree.”).
114. *Id.* at 167, 517 P.2d at 199; *see also* WASH. CONST. art. IV, § 1 (“Judicial Power, Where Vested. The judicial power of the state shall be vested in a supreme court, superior courts, justices of the peace, and such inferior courts as the legislature may provide.”); *id.* art. IV, § 30 (“Court of Appeals. (1) Authorization. In addition to the courts authorized in section 1 of this article, judicial power is vested in a court of appeals, which shall be established by statute.”).
power in the courts of this state. Then, based on authority of \textit{O'Connor}, the court proceeded from that constitutional grant of power to examine the “duties” imposed by that authority. This step was significant, because what followed is not spelled out in the constitution (state or federal) or in any statute. It is court-made law, based on the way the deciding court and past courts (including the \textit{O'Connor} court) defined their duties:

These duties include, among others, the fair and impartial administration of justice and the duty to see that justice is done in the cases that come before the court. The administration of justice demands that the doors of the judicial system be open to the indigent as well as to those who can afford to pay the costs of pursuing judicial relief.

In the \textit{O'Connor} case we stated . . . :

Were this court to hold that the Supreme Court has the power to waive prepayment of costs and that the superior court has a like power, but that no such power exists in justice courts, an anachronism would result. This would be tantamount to a holding that, if a poor person has a large claim, the courts will open their doors to him; but if his claim is small, those doors must be closed, simply because there were no justice courts at common law. . . .

The proper and impartial administration of justice requires that these doors be kept open to the poor as well as to those who can afford to pay the statutory fees. The court concluded from all of this that it had the constitutional duty to allow Lil Iverson to prosecute her appeal without cost.

Now, the important dimension that \textit{Iverson} adds to the access to justice struggle is the constitutional dimension. But it does so in a characteristically common law way. The state constitution says simply that judicial power is vested in the courts. But what does the court do? It decides what that judicial power entails. And it does so by looking at what it said in the past on the same question—what it said in \textit{O'Connor}. The court then adds another important constitutional justification for the notion of access to justice—that the courts are a separate branch of

\begin{itemize}
    \item \textit{Iverson}, 83 Wash. 2d at 167, 517 P.2d at 199.
    \item \textit{Id}.
    \item \textit{Id}.
    \item \textit{Id} at 167–68, 517 P.2d at 199.
\end{itemize}
government—the branch vested with judicial power. And, more importantly, the court then defines its obligations to the poor: “The administration of justice demands that the doors of the judicial system be open to the indigent as well as to those who can afford to pay the costs of pursuing judicial relief.”119

After *Iverson*, access to justice is now firmly rooted in a state constitutional grant of authority to the courts and the common law of this state.

C. Carter v. University of Washington

In *Carter v. University of Washington*,120 the University of Washington fired Sidney Carter because he violated regulations (which ones and how are unstated). He appealed to the Higher Education Personnel Board and later to the superior court. Both affirmed his dismissal. He then tried to appeal to the Washington State Supreme Court without paying a filing fee or posting a bond.121 It was this failure to pay the filing fee that stimulated the court’s discussion in this case.

Only three justices signed the lead opinion, which reached out most expansively for authority to support the notion of access to the courts. First, the lead opinion stated that the idea of access to justice has been around for a long time, maybe even two millennia: “Universal access to the courts is certainly not a novel concept in the annals of jurisprudence. Access to the courts was prized and protected by the Romans over 2,300 years ago.”122

Next, the lead opinion remarked that the problem of access had been noted, but movement toward easier access to the courts had been “less than impressive”.

119. *Id.* at 167, 517 P.2d at 199.


121. *Id.* at 392, 536 P.2d at 620.

122. *Id.* at 393, 536 P.2d at 620 (citing John MacArthur Maguire, *Poverty and Civil Litigation*, 36 HARV. L. REV. 361 (1923)). A number of other law review articles are used to bolster this claim:


123. *Id.* at 393 n.2, 536 P.2d at 620 n.2.
Fifty years ago, in 1924–[19]25, the American Bar Association’s Committee on Legal Aid Work drafted a model Poor Litigant’s Statute which provided, *inter alia*, that a poor litigant would be excused from giving security for costs and from payment of any fees. But the ABA’s model statute apparently has had only nominal influence in most jurisdictions in the development of poverty law.\(^{124}\)

Additionally, the lead opinion noted that the courts serve as society’s “complaint desk.”\(^{125}\) And in that capacity, the courts implement part of the social contract between society and individual citizens. Judicial resolution of private and public grievances promotes peace and avoids violent resolution of those disputes.\(^{126}\) Again, authority for this comes from a series of law review articles.\(^{127}\)

The lead opinion also stated that public policy alone cannot be the basis for access to the courts (why it cannot be the basis of a judicial decision is left unstated) and, specifically, waiver of fees and costs on appeal. The opinion turned, therefore, to more structured sources of common law doctrine—case precedent and the state constitution.\(^{128}\) It relied on *O’Connor* and *Iverson* for the conclusion that the courts have the inherent power to waive costs. The conclusion was bolstered by citation to a series of California cases and a law review article.\(^{129}\) It then turned to the Washington State Constitution. It is in this part of the opinion that two additional justices joined for a five-to-four majority.

The constitutional argument supporting the opinion came in steps. First, the court argued that access to the courts is a prerequisite to every other right—there are no rights if there is no right to enforce any right.

\(^{124}\) *Id.* (citing generally Lee Silverstein, *Waiver of Court Costs and Appointment of Counsel for Poor Persons in Civil Cases*, 2 VAL. U. L. REV. 21 (1967)).

\(^{125}\) *Id.*

\(^{126}\) *Id.* at 393–94, 536 P.2d at 620–21.


\(^{128}\) *Id.* at 394–96, 536 P.2d at 621–22.


275
The right of access must, therefore, be implied, like the right to travel and the right to vote. The concurring justices agreed only on those arguments grounded in the actual wording of two provisions of the Washington State Constitution: article I, section 4 (the right to petition) and article I, section 12 (requirement of equality of privileges and immunities).

Second, the *Carter* decision held, based on the right to petition, “that the explicit provision in our constitution preserving the right to petition for grievances encompasses and, indeed, makes fundamental the right of access to the courts.”

The *Carter* decision next grounded the right to access in Washington’s privileges and immunities clause. Before launching into this last bit of constitutional analysis, the court reached out for one more bit of custom to justify the use of the state constitution—Washington’s populist tradition. Its source for this bit of cultural information, which soon became law, is history.

And finally, the court reasoned that only a compelling state interest would justify denying indigent citizens access to the courts under the state privileges and immunities clause. The interest advanced by the state was the deterrence of frivolous lawsuits and controlling work load. But the court, relying yet again on a series of law review articles, concluded that this interest was constitutionally too broad, given other available safeguards.

At the end of the day, the only grounds for a right of access to justice that five justices could agree on were state constitutional grounds—the right to petition the government (article I, section 4) and the guarantee of equal privileges and immunities (article I, section 12). And, while the

130. *Id.* at 397–98, 536 P.2d at 622–23.
131. *Id.* at 403, 536 P.2d at 625.
133. “No law shall be passed granting to any citizen, class of citizens, or corporation other than municipal, privileges or immunities which upon the same terms shall not equally belong to all citizens, or corporations.” *Id.* art. I, § 12.
134. *Carter*, 85 Wash. 2d at 399, 536 P.2d at 623.
136. *Id.* at 400–01, 536 P.2d at 624.
137. *Id.* at 403, 536 P.2d at 625–26 (Utter, J., concurring).
most interesting and far reaching sources of authority are set out in the lead opinion, it is not the majority opinion. Only the Washington State Constitution provides the basis for the majority holding—a holding partially undone in Housing Authority v. Saylors.138

D. Housing Authority v. Saylors

The King County Housing Authority evicted Frances Saylors from her home for maintaining a “nuisance.” An administrative tribunal, composed of fellow tenants, affirmed the action, as did a superior court judge in the Housing Authority’s unlawful detainer action. Ms. Saylors, armed with the court’s decision in Carter, marched up to the Washington State Supreme Court and asked for public funds to prosecute her appeal.139 The court rejected her petition.140

The rationale of the Saylors decision is straightforward. The court interpreted the Washington State Constitution’s privileges and immunities clause (article I, section 12) similarly to the way in which federal courts have interpreted the equal protection clause of the Fourteenth Amendment.141 The United States Supreme Court, in construing the Fourteenth Amendment, does not require waiver of court fees for indigents when the right at stake is not a fundamental one and “there is another procedure available, not requiring the payment of fees, through which redress can be sought.”142 Therefore, the court held, since litigation in the field of economics and social welfare does not involve a suspect class, no special scrutiny is required.143

So, after Saylors, indigent civil litigants are back to relying on O’Connor and Iverson. To repeat the essential instructions of those cases: The courts have inherent power to waive fees and costs in civil appeals based on tradition and the constitutionally vested inherent authority of the courts if (1) the indigent litigant makes a showing of

139. Id. at 733–34, 557 P.2d at 322–23.
140. Id. at 744, 557 P.2d at 328.
141. Id. at 738–39, 557 P.2d at 325.
142. Id. at 739, 557 P.2d at 325.
143. Id. at 739–40, 557 P.2d at 325–26. On the question of whether poverty should be a suspect class, see Stephen Loffredo, Poverty, Democracy and Constitutional Law, 141 U. Pa. L. Rev. 1277, 1279 (1993), stating the obvious that the poor do not fully participate in the democratic processes and therefore democratic legislation disfavoring them should be reviewed with more scrutiny than simple rational basis.
poverty, (2) justice requires, and (3) the appeal is not frivolous and is prosecuted in good faith.

The dissent in Saylors reiterates the constitutional principles laid down in Iverson, Carter, and O’Connor. First, the Washington State Constitution (article IV, sections 1 and 30) vests the courts of this state with the power to administer justice (a principle not challenged by the majority): ""These duties include, among others, the fair and impartial administration of justice and the duty to see that justice is done in the cases that come before the court."" 144 Second, ""[t]he administration of justice demands that the doors of the judicial system be open to the indigent as well as to those who can afford to pay the costs of pursuing judicial relief."" 145 Third, the right to petition for redress of grievances, pursuant to article I, section 4, is not limited to cases involving political grievances. The court should ""construe article I, section 4, so as to give full effect to its broad language and purpose to the end that both the poor and the affluent may be heard by government."" 146 Finally, article I, section 12, of the state constitution must be interpreted differently than the Fourteenth Amendment to the United States Constitution. When properly applied, this provision is more protective of individual rights. 147

Recently, the constitutional rationale of Saylors (reliance on the United States Supreme Court’s view of the Fourteenth Amendment’s privileges and immunities clause) has been eroded by the decision in Grant County Fire Protection District No. 5 v. City of Moses Lake. 148 There, residents challenged the Washington state statutory scheme that permitted the petition method (by landowners) rather than an election method (by residents of the area to be annexed) as a means of annexing property to a city. Under the petition method, the petition can be based upon ownership of the majority of property in an area to be annexed.

144. Saylors, 87 Wash. 2d at 744, 557 P.2d at 328 (Horowitz, J., concurring in result only) (quoting Iverson v. Marine Bancorporation, 83 Wash. 2d 163, 167, 517 P.2d 197, 199 (1973)). This common law notion is all that remains as support for access to justice at the end of Saylors.
145. Id. at 744–45, 557 P.2d at 328 (Horowitz, J., concurring in result only) (quoting Iverson, 83 Wash. 2d at 167, 517 P.2d at 199).
146. Id. at 755–56, 557 P.2d at 334 (Horowitz, J., concurring in result only).
147. Id. at 756–57, 557 P.2d at 335 (Horowitz, J., concurring in result only).
148. 145 Wash. 2d 702, 42 P.3d 394 (2002), rev’d on rehe’g, ___ Wash. 2d ___, ___ P.3d ___ (2004) ("And, although in recent cases this court has held that the privileges and immunities clause is substantially similar to the equal protection clause, Seeley v. State, 132 Wn. 2d 776, 788, 940 P.2d 604 (1997), the possibility that article I, section 12 could be analyzed separately from the federal equal protection clause has been left open.").
rather than upon an equal vote by those residing in the area. The court engaged in a *State v. Gunwall* analysis and held:

> [T]he Gunwall factors weigh in favor of a determination that article I, section 12 of the Washington State Constitution provides greater protection than the equal protection clause of the United States Constitution when the threat is not of majoritarian tyranny but of a special benefit to a minority and when the issue concerns favoritism rather than discrimination.

The important point of *Grant County* for the access to justice community is that analysis of our own state’s privileges and immunities clause is not tied to federal Supreme Court Fourteenth Amendment jurisprudence. The constitutional support for access laid out in *Carter* and the dissent in *Saylors*, therefore, remains viable.

It is certainly true, as the *Saylors* decision notes, that democratically-elected legislators are appropriately vested with authority to legislate in the area of social and economic welfare. But it is also true that access to the court to redress grievances cannot be limited by either economics or social status. That remains a constitutional matter, the resolution of which is reserved to the courts.

### E. Doe v. Puget Sound Blood Center

*Doe v. Puget Sound Blood Center* is the last in this series of access to justice cases. I will first make a few observations about this decision. First, the author for this eight-judge majority also signed the *Saylors* opinion, which, as I have noted, unraveled some of the constitutional underpinning for access to justice. Next, the access issue was ancillary to the main question—whether the trial court abused its discretion by requiring production, subject to restrictions, of the name of a blood donor. The court nonetheless launched into a discussion of access and, in doing so, criticized the analysis of the *Saylors* decision.

---

149. *Id.* at 709–12, 42 P.3d at 397–99.

150. 106 Wash. 2d 54, 720 P.2d 808 (1986).

151. In *Gunwall*, the court settled upon six nonexclusive criteria to determine whether the Washington State Constitution extended broader rights to its citizens than the United States Constitution and should therefore apply, incidentally relying on a number of law review articles in the process. *Id.* at 59–61, 720 P.2d at 811–12.

152. *Grant County*, 145 Wash. 2d at 731, 42 P.3d at 408 (emphasis added).


Doe alleged that he contracted AIDS from a blood donation. He sued the Puget Sound Blood Center (Blood Center). Doe later died—the suit alleged as the result of AIDS—and the superior court substituted his estate as party plaintiff. The estate moved for an order requiring the Blood Center to disclose the identity of the donor of the infected blood. The Blood Center refused. The court exercised its discretion and required production of the identity subject to restrictions.\textsuperscript{155} The Washington State Supreme Court ultimately affirmed the court’s exercise of discretion.\textsuperscript{156} In doing so, however, it placed the discovery question in the broader context of access. In that broader context, the court made a number of noteworthy comments about access in general, and the \textit{Saylors} decision in particular. First, the court observed that “\textit{[o]ur cases on the right of access are somewhat perplexing.}”\textsuperscript{157} Next, the court quoted from the \textit{Iverson} discussion of access to justice: “\textit{[t]he administration of justice demands that the doors of the judicial system be open to the indigent as well as to those who can afford to pay the costs of pursuing judicial relief,” and \textit{[c]onsistent with our affirmative duty to keep the doors of justice open to all with what appears to be a meritorious claim for judicial relief, we hold that the plaintiff is entitled to the relief requested [waiver of fees and costs].}”\textsuperscript{158}

Commenting on the \textit{Saylors} court’s rejection of the \textit{Carter} court’s grounding of the right of access in sections 4 (right to petition) and 12 (privileges and immunities) of article I of the Washington State Constitution, the court noted: “The \textit{Saylors} court held that reliance upon the cited constitutional provisions was in error. \textit{H}owever, the important \textit{p}oint in \textit{Saylors} is the statement that \textit{[a]ccess to the courts is amply and expressly protected by other provisions.}’ Unfortunately, the court did not explore the rationale for its conclusion.”\textsuperscript{159}

Finally the court returned to a rationale expressed in \textit{O’Connor}, \textit{Carter}, and \textit{Iverson}: “\textit{The right of access is necessarily accompanied by those rights accorded litigants by statute, court rule or the inherent powers of the court, for example, service of process, [Revised Code of

\begin{footnotes}
\begin{enumerate}
\item Id. at 775–76, 819 P.2d at 372–73.
\item Id. at 776, 819 P.2d at 372.
\item Id. at 781, 819 P.2d at 375.
\item Id., 819 P.2d at 375 (emphasis omitted) (quoting \textit{Iverson}, 83 Wash. 2d at 167–68, 517 P.2d at 190).
\item Id. at 781–82, 819 P.2d at 375 (emphasis added) (citation omitted) (quoting Hous. Auth. v. \textit{Saylors}, 87 Wash. 2d 732, 742, 557 P.2d 321, 327 (1976)).
\end{enumerate}
\end{footnotes}
Common Law Process

Washington] 4.28, or statutes of limitation.”\textsuperscript{160} The court rejected as inadequate the Blood Center’s submissions in support of nondisclosure: “The Blood Center and allied amici attempt to support these contentions with materials ranging from a quote from \textit{Newsweek Magazine} to a quote from the Secretary of State of Scotland to testimony before a congressional subcommittee.”\textsuperscript{161} Among the court’s reasons for rejecting those submissions was the absence of a factual basis, either lay or expert.\textsuperscript{162}

V. ACCESS TO JUSTICE IN WASHINGTON—IN SUM

By way of summarizing this tour of Washington access to justice cases:

1. The right of an indigent person to sue is fundamental and significant.
2. Indigence does not mean complete destitution or total insolvency.
3. The authority for the courts to accommodate the petitions of indigent civil litigants is well grounded in English common law and English legislative enactments, both of which are part of our own common law.
4. “The right of access is necessarily accompanied by those rights accorded litigants by statute, court rule or the inherent powers of the court, for example, service of process, RCW 4.28, or statutes of limitation.”\textsuperscript{163}
5. The Washington State Constitution vests judicial power in the courts\textsuperscript{164} whereby the courts have assumed certain duties. These duties include the “fair and impartial administration of justice.”\textsuperscript{165}
6. The further constitutional grounding of access to justice in, at least, the privileges and immunities clause of article I, section 12, is still an open question.
7. And all of this is so because a common law court, relying on a variety of traditional and present day sources, has said so. Any

\textsuperscript{160} Id. at 782, 819 P.2d at 375.
\textsuperscript{161} Id. at 786, 819 P.2d at 377.
\textsuperscript{162} Id. at 786–87, 819 P.2d at 377–78.
\textsuperscript{163} Id. at 782, 819 P.2d at 375.
\textsuperscript{164} \textsc{Wash. Const.} art. IV, §§ 1, 30.
accommodation to the poor through the implementation of the ATJ-TBoR principles is necessarily premised on the recognition of a fundamental right of access to justice. And that right is the product of common law judges’ application of traditional common law processes to the resolution of individual cases.

A final thought: Unrestricted access by the poor to our legal system for the resolution of their disputes is important in one other way. It has implications for the development of law. Common law courts articulate and bring to life the values of a society. They do so in the form of legal principles—principles developed in case law. Every single appellate case, then, adds to the body of law. If our guiding principles are truly to reflect the full range of the values of our society, and not just a narrow segment of it, all members of the community must have at least the opportunity to influence their development. So, to the extent that the poor are denied access to the courts, the principles articulated by the courts will not reflect the general customs and traditions of this state.

Two basic questions repeatedly crop up in the judicial debate over access to justice. First, is it a good idea? This is a question posed by those who tend toward the more economic model of justice and rights. Their concern is that easy access simply opens up the floodgates of litigation, including lots of frivolous lawsuits that economic constraints now hold in check. In this view, access is a public policy and, therefore, solely a legislative concern. The second recurrent question is, if it is a good idea, then is it required to be implemented, or at least accommodated, by our state constitution? Of course, if access to justice is constitutionally required, then the first question goes away. It makes no difference whether access to justice is a good idea.

These ATJ-TBoR principles reflect economic, philosophical, and moral judgment; they all sound like a good idea: that is, as a purely rational proposition, as simply a matter of scholarly deduction. Their justification will be developed on a case-by-case basis by the courts.

VI. CONCLUSION

The proposed principles memorialized in the ATJ-TBoR fit into our modern common law system. It is impossible to divine the ultimate effect of every new technology on every rule, practice, or procedure. But


what we can do in the common law tradition is promulgate a set of principles that public rulemaking bodies—the legislature, the courts, the American Law Institute—can use so as to avoid the possibility, indeed the probability, that technology will reduce or inhibit the enjoyment of a fundamental right—the right of access to the courts.168

The best that any rulemaking body can do is craft a set of general principles. Those principles will then influence the exercise of discretion in the courts in a way that improves access to the courts in spite of and, we hope, because of technology. The specifics, then, for implementing this general set of principles for a technology bill of rights are best left to judges, passing upon individual cases brought about by particular controversies that come before the court.

By advancing the common law system, I do not suggest that we turn back the clock to the days when common law judges were regarded as oracles.169 I look instead to the ways in which a common law system functions in this state at this moment in our history. The judicial approach to legislation is an imperfect, but helpful, parallel. The court reads legislation so as to effect the purpose of the legislative body.170 Likewise, other rules are applied by the courts on a case-by-case basis to effect their articulated purpose—rules of evidence, rules of practice and procedure, and statutory enactments affecting the processing and filing of litigation.171

The modern common law judge usually works with rules promulgated by someone else—the legislature, administrative agencies, state supreme courts in their rulemaking capacity, or uniform law committees. But

---

168. One concern, for example, is e-filing of court documents by proprietary systems (such as West or Lexis) upon which users will become as dependent for electronically filed documents much as they presently are for judicial opinions. See, e.g., Wendy R. Liebowitz, Courts Electrify Suits, Sparks Fly: New Rules Needed for E-filings, NAT'L LAW J., Sept. 7, 1998, at B6, cited in Dale A. Whitman, Digital Recording of Real Estate Conveyances, 32 J. MARSHALL L. REV. 227, 268 (1999). Off-line members of the public will need assistance to access electronic court files. Lederer, supra note 15, at 804–05.


170. See Cockle v. Dep’t of Labor & Indus., 142 Wash. 2d 801, 807, 16 P.3d 583, 585–86 (2001) (“The primary goal of statutory construction is to carry out legislative intent.”); State v. Fjermestad, 114 Wash. 2d 828, 837, 791 P.2d 897, 902 (1990) (Guy, J., dissenting) (“Under general principles of statutory construction, when construing a statute the court’s purpose is to ascertain and give effect to the intent of the Legislature.”).

whatever the source, it is the application of these rules that must reflect the needs of a particular case, and it is that case that will be cited to by future courts when the same or a similar question is presented. It is in that way that law becomes “self-generating.” 172 Another benefit of a common law system is its ability to adapt to changing conditions 173 — “to meet these new and unexpected conditions of society.” 174 Witness the changes in the law that accommodated the industrial revolution or the current changes with the scientific advances in DNA.

The common law has the capacity to adapt to new technologies. Of concern here is the way in which it adapts. Ultimately, judges must sit and reason to the notion that all of our fellow men and women should have equal access to these courts. That is what we have tried to do with the ATJ-TBoR.

Is it perfect? Of course, it is not. But it is a start. It provides the judges of this state with the ability to craft rulings that at a minimum do not impede access to the justice system and perhaps may enhance the access of the poor to the system. Judges must reconcile often inconsistent and contradictory values. 175 Adoption of the ATJ-TBoR in some form is in keeping with a long tradition of not only access to the courts by the poor but also liberal resort to extrajudicial sources for principles to effect such rights. These principles should, by a process of accretion, become a part of the common law here in Washington. As such, they will be part of a principled approach to adapting to the state’s changing social, cultural, and political conditions. 176

172. See HANKS, supra note 6, at 4.
173. Madden, supra note 43, at 593.
CRAFTING A LICENSE TO KNOW FROM A PRIVILEGE TO ACCESS

Jane K. Winn

Should the doctrine of trespass to chattels apply to unauthorized access to Internet facilities? If it does, then the property rights of the owners of computers connected to the Internet may be vindicated, but at a cost of diminished public access to information posted on the Internet. If it does not, then incentives to invest in the kind of commercial facilities that now largely constitute the Internet may be undermined, but the public interest in knowledge gleaned from information posted on the Internet will be protected. Although trespass to chattels has been derided as an anachronism ill-suited to the Internet, and its application to Internet activities rejected in some recent cases, other cases have held decisively that its application gives appropriate recognition to the rights of owners of computer equipment connected to the Internet. In order to

* Director and Professor, Shidler Center for Law, Commerce & Technology, University of Washington School of Law. Thanks to William Edmundson, Brad Handler, and Jay Monahan for helpful comments.

1. Trespass to chattels is defined as the unauthorized, intentional, and substantial use of or intermeddling with another’s tangible personal property. RESTATEMENT (SECOND) OF TORTS §§ 217–218 (1965).


safeguard the “license to know” factual information posted to the Internet that the public currently enjoys, courts should recognize an individual privilege to access Internet resources in a reasonable manner.

Given that trespass to chattels is unlikely to disappear from the Internet landscape any time soon, refinements are needed to keep the doctrine’s scope within reasonable bounds and to make its application more predictable. The California Supreme Court recently imposed such a limitation on its application by holding that liability for trespass should be found only if the Internet access at issue significantly impairs the functions of another’s computer equipment or, if widely replicated, would so impair it.\(^5\) However, this attempt to restrict the scope of earlier rulings may prove to be at least as contentious as the holdings of the cases it purports to limit, and so is unlikely to staunch the flow of controversy.

The California Supreme Court focused on the functional impact that unauthorized access has on computer equipment owned by the party objecting to the access. A more helpful refinement of trespass doctrine might be found by considering instead which forms of access equipment owners have consented to merely as a consequence of connecting their equipment to the Internet. In every case in which trespass to chattels has been raised as an issue, before filing suit the equipment owner had demanded in no uncertain terms that the unauthorized access stop immediately, so the accessing party obviously cannot rely on a defense of express or implied consent. Courts could instead recognize a form of “constructive” consent to certain reasonable forms of access\(^6\) that would defeat a claim of trespass. While such a finding of “consent” would not correspond to the actual subjective state of mind of the plaintiff bringing a trespass to chattels claim, it would have the benefit of refocusing attention on the social significance of the public character of the Internet, and hold the owner of computer equipment connected to the Internet accountable for having made the choice to create that connection.

---

\(^5\) Hamidi, 71 P.3d at 296, 306.

Crafting a License

One of the most significant problems created by the application of trespass to chattels doctrine to unauthorized Internet access disputes is its overbreadth. Trespass doctrine lacks the nuances normally found in intellectual property law to balance competing public and private interests in the exploitation of ideas and knowledge. Overbroad grants of rights in information have a chilling effect on the progress of science and the dissemination of knowledge generally.\(^7\) Trespass doctrine vindicates the property rights of equipment owners at the expense of the “ease and openness of communication”\(^8\) that has always been the hallmark of the Internet. Overzealous application of trespass doctrine obscures the fact that some forms of Internet access must be privileged if the unique public character of the Internet is to be preserved. Such a privilege should be limited in scope in recognition of the role played by private parties in maintaining the Internet today. The recognition of this privilege, however, should not be made contingent on the voluntary acquiescence of private parties.

Recognizing a defense to a claim of trespass in Internet cases based on a finding of constructive consent provides a doctrinal basis for privileging some forms of access while acknowledging a right to exclude certain other forms of access. Focusing attention on the public character of the Internet and assigning a clear legal significance to the equipment owner’s deliberate choice to participate in that arena provide a more secure legal foundation for such a privilege to access than the “functional impairment” standard offered by the California Supreme Court. The contours of such a doctrine of constructive consent to Internet access are suggested by the terms of the license eBay offered to Bidder’s Edge as discussed below—access by individual Internet users or its functional equivalent. This Article suggests that a defense based on constructive consent can complement the limitation imposed by the California Supreme Court to further limit the scope of trespass doctrine in Internet arenas, increase the predictability of the doctrine’s application in new disputes, and help to protect important public interests in free and open access to Internet resources.

\(^7\) See, e.g., J.H. Reichman & Pamela Samuelson, *Intellectual Property Rights in Data?*, 50 VAND. L. REV. 51 (1997) (arguing that an intellectual property right in data will be harmful to science and education if based on the labor expended in building the database instead of innovation contained in the data).

\(^8\) Hamidi, 71 P.3d at 311.
I. SETTING THE STAGE: EBAY, INC. V. BIDDER'S EDGE, INC.

The first case to apply the doctrine of trespass to chattels to modern networked communication services involved the unauthorized use of access codes to make long distance calls. That case held that teenagers who hacked into a long-distance telephone service’s computer system and made unauthorized long distance calls using access codes thus obtained could be held liable for trespass to chattels. The chattels at issue were the telephone access codes; the court declined to limit the application of trespass to chattels doctrine to intangible interests that were clearly associated with an interest in tangible property.

The second case applying trespass to chattels to a modern networked communication system involved Cyber Promotions using the facilities of CompuServe, an Internet service provider (ISP), to send unsolicited commercial email (also known as spam) to the ISP’s subscribers. CompuServe’s subscribers threatened to terminate their subscriptions unless it could stop Cyber Promotions from spamming them, and CompuServe undertook every technological measure at its disposal in a fruitless attempt to block Cyber Promotions’ communications. The court found that Cyber Promotions’ spamming constituted trespass to chattels, the chattels in question being CompuServe’s computer equipment, because the activity deprived CompuServe of the economic value of the equipment even though it did not lose possession of it. The court rejected the argument that Cyber Promotions’ access was privileged because CompuServe had consented to receive spam addressed to its subscribers, finding instead that whatever consent might be inferred from connecting its equipment to the Internet had been revoked by communications from CompuServe to Cyber Promotions.

Perhaps the most well-known case to apply the theory of trespass to chattels involved a conflict between eBay and Bidder’s Edge. eBay’s primary business involves providing an Internet auction service that permits individuals to offer items for sale and also to purchase items

---

10. Id. at 473.
11. Id. at 473 n.6.
13. Id. at 1019.
14. Id. at 1022.
15. Id. at 1024.
Crafting a License

posted for sale. eBay’s business model focuses on creating a “community” of buyers and sellers; eBay does not itself offer for sale or guarantee any of the products offered on its site. Amazon.com and Yahoo! have similar auction sites, but eBay dominates the U.S. market for Internet auction services. These Internet auction sites are designed to give unrestricted access to individuals wishing to browse the items offered for sale. To place an item for sale, or to bid on an item, it is necessary to register with the auction site, and the registration process requires individuals to manifest their assent to the auction site’s terms and conditions of use.

A. Auction Aggregators: Licensed and Unlicensed

During the exuberant days of the dot com bubble, public interest in Internet auction sites such as eBay was exploding. A handful of dot com entrepreneurs devised the “auction aggregator” business model to permit individuals to search more than one Internet auction site simultaneously and to learn at once which site offered the best deal. There are no auction aggregators in operation today, but aggregators that once captured quite a bit of attention included www.auctionwatch.com, www.auctionferret.com, and www.auctionrover.com.

Given that eBay was the dominant player in the U.S. Internet auction site then as well as now, it could be expected to have ambivalent feelings about aggregator sites. Aggregators might increase the liquidity of auction markets created by its competitors; however, to the extent that eBay offered the largest selection and best prices, comparisons might have only increased its advantage over its competitors. To the extent that eBay’s sellers would gain access to a larger group of prospective buyers through referrals from aggregators, however, they could be expected to support the work of the aggregator sites.

Perhaps in order to accommodate the wishes of its sellers, eBay made a practice of licensing to aggregators access to information about auctions taking place on eBay’s site. These licenses were granted subject to certain restrictions designed to minimize the demands placed on eBay’s own system by the aggregators, and to guarantee the accuracy


of information being provided to aggregators’ customers. Aggregators were authorized to provide comparisons to aggregator site end users, provided that the demands they placed on eBay’s site were equivalent to the demands that would have been placed on its systems if aggregators’ customers had visited eBay’s site directly. In other words, eBay would only authorize aggregators to perform “real time proxy searches,” in which a query would be submitted for one item at a time, and the current price for that item would be provided to the aggregator’s end user immediately. This type of proxy search both limited the demands that aggregator sites placed on eBay’s servers and guaranteed that the aggregator’s end user was provided with updated, accurate price information.

Many aggregators were willing to live within these constraints, but some, including Bidder’s Edge, were not. Bidder’s Edge’s business model involved sending software robots (bots) onto eBay’s site once every twenty-four hours to copy information about everything offered on the site. This information was sent back to Bidder’s Edge and displayed in response to Bidder’s Edge’s end users’ queries until another copy of all the listing information on eBay was made the following day. eBay objected both to the demands that the bots were placing on its system when copying all the listing information at once, and to the fact that Bidder’s Edge’s end users were often being shown inaccurate price information. When Bidder’s Edge’s end users clicked through an eBay listing only to find that the price had gone up since the Bidder’s Edge copy of the data had been made, some of them blamed eBay for the unexpected change in price. eBay was also concerned that Bidder’s Edge’s wholesale approach to collecting data off its site might cause eBay to fail to perform some of its undertakings under its privacy policy, and might make it possible for Bidder’s Edge to reuse information about “members” of the eBay community in ways that were expressly prohibited under the eBay User Agreement.

18. E-mail from Jay Monahan, Vice President, eBay, Inc., to author (Sept. 23, 2003) (on file with author).
Crafting a License

eBay signaled its unwillingness to permit bots to trawl its site by using “robot exclusion headers.” Many Internet businesses that rely on information collected by bots to function, such as search engines, program their bots to abide by restrictions placed by website operators in robot exclusion headers. Bidder’s Edge, by contrast, decided that it could ignore the content of eBay’s robot exclusion headers with impunity because the information it was collecting from eBay’s site was simply factual, and thus unlikely to be protected by copyright. In addition, because eBay’s business model dictated that as much information as possible should be made publicly accessible on its site, and only participation in actual purchases and sales should require registration and the formation of a contract with eBay, Bidder’s Edge could argue that the eBay User Agreement also did not apply.

eBay and Bidder’s Edge entered into license negotiations that would have granted Bidder’s Edge permission to perform real time proxy searches on behalf of its end users, but the negotiations were broken off without agreement. eBay then began to use all technological means available to it at the time to block Bidder’s Edge’s bots from accessing its site, but without success. eBay next filed suit against Bidder’s Edge, seeking an injunction to prevent Bidder’s Edge from sending unauthorized bots onto eBay’s servers. The suit was based on various theories including trespass to chattels, false advertising, federal and state trademark dilution, computer fraud and abuse, unfair competition, misappropriation, interference with prospective economic advantage, and unjust enrichment. eBay was granted the injunction based on the trespass to chattels argument, but the court did not reach the other claims. Bidder’s Edge appealed to the U.S. Court of Appeals for the Ninth Circuit, but shut down its website before the court heard oral arguments. Shortly thereafter, Bidder’s Edge paid eBay an undisclosed sum to settle the litigation.

While the outcome in the eBay, Inc. v. Bidder’s Edge, Inc. case may have resolved the conflict between those two parties, the district court’s ruling established an overbroad precedent. The district court’s opinion does not place any limits on eBay’s power to restrict access to its site

23. Id.
that eBay offered to Bidder’s Edge in the failed license negotiations. eBay was willing to grant Bidder’s Edge access to the information its customers wanted in a way that guaranteed that Bidder’s Edge’s customers were provided with only accurate price comparisons, but Bidder’s Edge rejected the offer because it thought it could do better without a license.

B. Trespass to Chattels: Costs and Benefits

Applying the doctrine of trespass to chattels to the problem of unauthorized access to Internet resources has several benefits. It recognizes that most of the facilities that make up the Internet are now owned and operated by private parties. These are not eleemosynary institutions, and it is not their intention to donate their computers to the public once they connect them to the Internet. The doctrine of trespass to chattels also can be invoked if someone merely “intermeddles” with a chattel, a concept that is sufficiently vague to be capable of expanding to cover access to Internet resources. It puts the owner of the computer equipment in the driver’s seat with regard to determining what access to its equipment is acceptable and what is not.\(^\text{25}\) In the absence of such a theory, owners of computer equipment would have to consider the consequences of connecting their equipment to the Internet much more carefully, and, in all likelihood, some businesses would make the decision to withdraw their systems from full participation in the Internet in order to maintain an acceptable level of control over their networks.

These benefits notwithstanding, trespass to chattels fails to provide an adequate mechanism to balance the competing claims of Internet resource providers and Internet resource users. Asking whether an unauthorized electronic access to data stored in digital form on a server is equivalent to an unauthorized use of a toothbrush does not provide a rational basis for the development of the law in this area.\(^\text{26}\) The doctrine of trespass to chattels was considered archaic and underdeveloped before

\(^{25}\) See Epstein, supra note 4.

\(^{26}\) The Restatement (Second) of Torts comments:

There may, however, be situations in which the value to the owner of a particular type of chattel may be impaired by dealing with it in a manner that does not affect its physical condition. Thus, the use of a toothbrush by someone else may lead a person of ordinary sensibilities to regard the article as utterly incapable of further use by him, and the wearing of an intimate article of clothing may reasonably destroy its value in his eyes. In such a case, the intermeddling is actionable even though the physical condition of the chattel is not impaired.

RESTATEMENT (SECOND) OF TORTS § 218 cmt. h (1965).
Crafting a License

its sudden burst of fame in the Internet context, and its application to sophisticated computer technology is unlikely to produce profound insights into the character of social relationships mediated by technology and law. Unlike nuisance doctrine, it does not require the explicit balancing of the competing public and private interests affected by the regulation of Internet access, substituting a private, commercial decision-making process for a more public, participatory process characteristic of the Internet in its early days.\textsuperscript{27}

Perhaps the most appropriate theory for granting eBay the relief it sought against Bidder’s Edge, and in similar cases, might have been some kind of a reverse passing off\textsuperscript{28} “cold news” variation of the “hot news” misappropriation doctrine established in \textit{International News Service v. Associated Press}.\textsuperscript{29} In \textit{International News}, the U.S. Supreme Court held that a rival news service could not copy news stories and resell the information even though the news stories lacked copyright protection.\textsuperscript{30} The precise contours of misappropriation are somewhat unclear, but it seems at a minimum to grant a “quasi-property right” to the party claiming misappropriation, and to recognize the value of time and effort spent creating something of economic value that is not recognized by existing intellectual property law doctrines.

Although the holding in that case has been limited to its facts by subsequent cases,\textsuperscript{31} those facts share a common characteristic with the facts of the \textit{eBay, Inc. v. Bidder’s Edge, Inc.} dispute because a large part of the economic value of the commercial information in both cases was determined by its “freshness.” In the \textit{eBay} case, however, Bidder’s Edge’s behavior might have driven eBay’s customers away, not because they could get product of equal value from Bidder’s Edge, but because the Bidder’s Edge business model involved serving up eBay prices over eBay’s objection after they had become stale. Such a tenuous argument

\textsuperscript{27} See Burk, supra note 2.
\textsuperscript{28} Reverse passing off is the marketing of another’s product under a claim that it is one’s own; passing off is marketing of one’s own product under another’s mark, i.e., trademark infringement. See, e.g., Dastar Corp. v. Twentieth Century Fox Film Corp., 539 U.S. 23, __ n.1, 123 S. Ct. 2041, 2045 n.1 (2003).
\textsuperscript{29} 248 U.S. 215 (1918).
\textsuperscript{30} Id. at 219.
would have been unlikely to support a claim for a preliminary injunction, however, leading eBay to emphasize its trespass claim as a more solid basis for a grant of an injunction.

If the case had been decided on a misappropriation theory instead of a trespass theory, the court might have articulated limits to eBay’s discretion in deciding what type of access to grant visitors to its site. For example, the doctrine of misappropriation prohibits competitors from reusing information, but does not prohibit individuals in the general public from reusing the same information. 32 A holding based on a misappropriation or deceptive trade practices theory might have distinguished between prohibiting certain forms of access arising out of unfair competition and preserving open access for other Internet users not engaged in any form of unfair competition.

The controversy surrounding the appropriateness of applying trespass doctrine to Internet access disputes is unlikely to subside any time soon, nor is the inconsistency in the manner in which the doctrine has been applied likely to be eliminated soon. While it remains possible that Congress will act to resolve this turmoil by enacting legislation that balances the competing public and private interests fairly, it is unlikely that will happen in the near future. So both Internet site operators and visitors will likely be left struggling to make sense of the emerging jurisprudence of trespass to Internet facilities. Articulating more clearly the significance of the decision by the owner of computer equipment to connect it to the Internet may create a mechanism for establishing a better balance of public and private interests, thus diffusing some of the current controversy and providing greater predictability in the application of trespass doctrine.

II. A RIGHT TO EXCLUDE QUALIFIED BY A PRIVILEGE TO ACCESS

The debate over whether trespass to chattels should be applied to resolve disputes involving unauthorized Internet access grows more acrimonious with passing time. On the one side are the “propertization” advocates, arguing that property rights of owners of the computer equipment at issue should trump other interests, giving the property owners a unilateral right to veto any use of their equipment they do not like. 33 On the other side are the supporters of the idea of the Internet as

33. See Epstein, supra note 4.
Crafting a License

an open, public space where the community interest in preserving that openness conditions the right of owners of computer equipment to connect to the Internet on their acceptance of pre-existing Internet social norms of openness.34

The degree to which the debate has become polarized is obvious from this comment by Justice Brown in her dissent in Intel Corp. v. Hamidi35: “Those who have contempt for grubby commerce and reverence for the rarified heights of intellectual discourse may applaud today’s decision, but even the flow of ideas will be curtailed if the right to exclude is denied.”36 But the debate need not be so polarized. The attempt to frame each position in absolute terms distorts each argument and obscures a possible middle ground where the competing claims of equipment owner and Internet end user might be harmonized.

Recasting the arguments using Hohfeldian terminology of rights can help clarify this middle ground.37 Hohfeld suggested his system for classifying different forms of legal relations to show when apparent conflicts among different legal interests were misleading.38 In the case of unauthorized access to Internet sites, vindicating the property rights of equipment owners negates any possible right Internet site visitors might have to free and open access to information posted on the Internet. Vindicating the public interest in freely making use of information posted on the Internet negates any right of the equipment owner to exclude others from its equipment. While each side would like to claim a strong form of rights in support of its position, it may be more accurate to say that equipment owners have certain limited property rights bundled together with certain privileges and powers, while the end users clearly could also be found to have a privilege of access that the equipment owner must respect. Recasting the debate in these terms is merely a first step toward resolving the controversy, because even if equipment owners are prepared to concede that end users enjoy an

34. See Burk, supra note 2; Ruth L. Okediji, Trading Posts in Cyberspace: Information Markets and the Construction of Proprietary Rights, 44 B.C. L. REV. 545 (2003); O’Rourke, Shaping Competition, supra note 2.
35. 71 P.3d 296, 325 (Cal. 2003) (J. Brown, dissenting).
36. Id. at 325 (J. Brown, dissenting).
38. Hohfeld, supra note 37, at 18.
implied license to access public Internet sites, the scope of that license remains to be defined.

A. Right to Exclude

Property rights in tangible computer equipment should not be conflated with a much broader right to exclude Internet users from accessing information on public web sites. If web site operators have a right to exclude end users, then end users have a corresponding duty not to interfere with the web site operator’s exercise of its right. This is because, using Hohfeld’s taxonomy, “duty” is the jural correlative of “right.” But assigning strong rights and correlative strict duties in this manner is at odds with the reality of Internet use by both web site operators and end users. Web site operators connect their systems to the Internet precisely in order to avail themselves of the public character of the network for commercial advantage. If web site operators are concerned about controlling access to their equipment, then such control can be accomplished by technological means—albeit at a cost of reduced traffic to a site. For example, eBay permits casual visitors to look at auctions without registering, but will only permit registered “members of its community” to actually participate in actions as buyers or sellers. In order to join the community, individuals are required to complete a series of web forms and click through a contracting interface, agreeing to be bound by eBay’s User Agreement. Registered users have user IDs and passwords that they must use to participate in auctions. User IDs and passwords with little or no verification of the information provided is a very rudimentary form of access control. If eBay required more security, it could use other networking technologies such as “virtual private networks.”

When commercial parties choose to connect their computer equipment to the Internet without restricting access to that equipment through the use of technological access controls, they are choosing to participate in a public forum. The Internet’s public character was

39. See EDMUNDSON, supra note 37, at 154; Hohfeld, supra note 37, at 30.
40. Orin S. Kerr, Cybercrime’s Scope Interpreting “Access” and “Authorization” in Computer Misuse Statutes, 78 N.Y.U. L. REV. 1596 (2003) (statutes criminalizing unauthorized computer access should be interpreted as requiring the circumvention of a technological barrier to access rather than violating a contractual limitation on access).
established long before commercial exploitation of the Internet was permitted. Prior to 1995, the National Science Foundation Acceptable Use Policy (AUP) applied to Internet activity, and prohibited commercial use unless the National Science Foundation (NSF) reviewed the use for consistency with its overall mission and granted permission for it.42 At that time, the purpose of the Internet was “to support research and education in and among academic institutions in the U.S. by providing access to unique resources and the opportunity for collaborative work.”43 As Kesan and Shah explain:

The NSFNET (later known as the Internet) connected universities, federal agencies, public and private research laboratories, and community networks. While the NSFNET encouraged such diversity, it also had an Acceptable Use Policy (AUP). The AUP prohibited the use of the NSFNET for purposes not in support of research and education, a policy consistent with the NSF’s mission. Nevertheless, a growing number of users wished to use NSFNET for purposes beyond research and education, a push for what the NSF termed “commercial use.” The potential for commercial use of the Internet propelled regional networks to create for-profit spin-offs. These for-profit commercial networks would eventually form the basis for the privatized Internet backbone.44

So the first commercial uses of the Internet were possible because the NSF granted immunity from expulsion to for-profit entities that joined the Internet. Under that immunity, the number of for-profit service providers grew until 1995, when the NSF was able to withdraw its support for the Internet backbone and turn it over to private parties to operate.

Equipment owners that once accepted a mere immunity from expulsion in order to share in the benefits of Internet access are now trying to turn the tables on other Internet users and claim a right to exclude other Internet users at will. But this is too broad a claim of right:

44. Id. at 111 (citing BRIAN KAHIN, The NREN as Information Market: Dynamics of Public, Private, and Academic Publishing, in BUILDING INFORMATION INFRASTRUCTURE 323–24 (Brian Kahin ed., 1993)).
a narrower claim of right that recognizes a privilege of access for ordinary end users is adequate to protect the equipment owner from unfair competition and hostile intrusions to their equipment that interfere with their commercial activities. A narrower claim of right more fairly balances the interests of equipment owners and ordinary end users by permitting web site operators only to restrict access by other commercial parties, not access by individual end users. Some form of implied license for individual end users must be recognized to prevent the destruction of the open, public character of the Internet in the name of commerce.

Granting web site operators a power to exclude other commercial actors from overbroad access to their sites, in lieu of a stronger right to exclude anyone from accessing their sites for any reason, assigns a realistic and appropriate significance to the equipment owner’s free choice to connect its equipment to an open, public communications medium. In Hohfeldian terms, to say that someone has a power is not to say that anyone else has a duty; rather, someone else might incur a duty if the power is exercised.45 Giving web site operators a power to fend off potential competitors permits them sufficient control over their equipment to protect its commercial value without depriving the general public of its ability to enjoy freely the open character of the Internet. If web site operators want to restrict access to their sites by the general public, then they can take concrete steps to restrict access to information on servers attached to the Internet, for example, by putting the information behind a firewall and implementing technological access controls. Many commercial web site operators, such as eBay, are unwilling to place these kinds of restrictions on casual visitors to its site, but do place these restrictions on anyone who would proceed from merely viewing information to transaction processing. The business decision regarding the design and implementation of access controls to Internet facilities should be assigned a legal significance in any subsequent dispute over whether a particular form of access was authorized.

B. Right to Access

Even the most vigorous opponents of the application of trespass to chattels to the issue of Internet access have not argued that Internet end

45. See EDMUNDSON, supra note 37, at 155.
Crafting a License

users have a “right to access” Internet facilities. Such a strong claim would imply that equipment owners have a duty to maintain the equipment so that end users’ rights can be exercised. Instead, opponents argue that end users should receive a broad grant of immunity from liability as a consequence of the equipment owner having made the decision to connect to the Internet. In the eBay case, this immunity would have prevented eBay from getting an injunction to stop Bidder’s Edge’s bots from copying and transmitting large quantities of information accessible on eBay’s site before the court reached a judgment on the merits.

A commercial entity such as Bidder’s Edge cannot claim that the type of access for which it demands immunity was an integral part of the public character of the Internet that eBay knowingly embraced. As early as 1994, a standard for robot exclusion was being developed informally to permit web site operators to communicate their desire to exclude software robots from their sites. Bidder’s Edge’s decision to send bots to copy and transmit data from eBay’s site for commercial exploitation on Bidder’s Edge’s site bears no resemblance to the types of access that would have been permitted under the NSF’s AUP. By contrast, the terms of the license that eBay offered Bidder’s Edge and that Bidder’s Edge rejected bore a close resemblance to the types of access that would have been permitted under the NSF’s AUP. Bidder’s Edge rejected eBay’s “reasonable access” license, however, and instead gambled on an aggressive claim that because it had the right to make unrestricted commercial use of the factual information on eBay’s computers, eBay had no right to restrict its access to eBay’s servers.

C. Privilege to Access from Constructive Consent

Consent may create a defense to a claim of tort liability by creating a privilege to engage in the conduct in question. An end user has a plausible claim that any web site operator who has not articulated an

47. “[Hamidi] does not argue that he has a right to force unwanted messages on Intel.” Intel Corp. v. Hamidi, 71 P.3d 296, 318 (Cal. 2003).
49. RESTATEMENT (SECOND) OF TORTS § 890 cmt. b (1965).
express policy governing access to its site has impliedly consented to any reasonable, conventional form of Internet access. However, this implied license alone cannot support the creation of a robust privilege of access for individual end users because a website operator can revoke any implied consent at any time by notifying visitors that it has established a restrictive access policy.

The only way to create a robust privilege of access for the general public is to find constructive consent to access websites based on a website operator’s choice to join the Internet without placing any functional restrictions on access to its site. Constructive consent is not a finding that consent exists as a factual matter, but rather is a legal fiction that asserts that something tantamount to consent does exist, and that it will be given the same legal effect as consent. While courts are often willing to find consent implied in light of parties’ behavior in a given context, they are generally reluctant to invoke the notion of constructive consent without an extraordinary justification.50 This general reluctance notwithstanding, a finding of constructive consent can be used to balance competing public policy objectives.51 Here, the competing policy objectives are the need to allow website operators to protect themselves against interference, and the public’s need for open access to websites. The notion of constructive consent shifts the obligation from the individual end user to ensure that his or her access is permitted to the website operator to choose between granting the general public reasonable but unfettered access to its site and placing some form of functional access controls on its site.

In order to strike an appropriate balance between the interests of individual end users in preserving the public character of the Internet and the interests of commercial website operators in preserving and exploiting the value of the equipment they have connected to the Internet, the notion of constructive consent must be limited to those situations where the public interest in unrestricted access is clear. The NSF’s AUP provides a convenient starting point for the process of defining what “reasonable” individual access or its functional equivalent would be. But because AUP ceased to apply to Internet activities nearly a decade ago, it would be anachronistic to adhere too closely to its terms.

50. See, e.g., In re Pharmatrak, Inc. Privacy Litig., 329 F.3d 9, 19–20 (1st Cir. 2003); Griggs-Ryan v. Smith, 904 F.2d 112, 117 (1st Cir. 1990).
Crafting a License

in the search for a definition of access to Internet facilities so reasonable, it should always be privileged.

Requiring a finding that machine functions were neither actually impaired nor likely to be impaired as a result of wide replication of the offending form of access, as the California Supreme Court did in *Hamidi*, is an important element in recognizing a privilege to access public Internet facilities, but by itself is not a complete protection for that privilege. The focus of a court’s analysis should be on the intent of commercial operators of computer equipment in making the decision to connect their facilities to the Internet, not on the impact that access has on machine functions. Focusing on intent and the social significance of the public character of the Internet creates a framework within which a privilege to access Internet resources can evolve with technological change while remaining consistent to its objectives.

III. RECENT TRESPASS CASES

The holdings of recent cases applying the doctrine of trespass to chattels to Internet access disputes veer from finding no privilege to access Internet sites to finding immunity from liability for clearly unauthorized access. If Internet facility operators, by connecting their equipment to the Internet without technological access controls, are deemed to have consented to the technological equivalent of individual access, can this notion of constructive consent help to make sense of recent trespass cases that otherwise seem to veer back and forth between contradictory interpretations? As the following analysis makes clear, the privilege of reasonable access for individual Internet users or its functional equivalent unfortunately is not a silver bullet that magically resolves all the overbreadth problems inherent in applying trespass to Internet access disputes. It may nevertheless help to focus attention on which characteristics of the Internet commercial parties should be required to tolerate, however grudgingly, as a condition of maintaining an open connection between their equipment and the Internet.

A. Register.com, Inc. v. Verio, Inc.

In Register.com, Inc. v. Verio, Inc., a domain name registrar used trespass to chattels to stop unauthorized access to its site by a competitor. Although Register.com was required by its Registrar Accreditation Agreement (RAA) with the Internet Corporation for Assigned Names and Numbers (ICANN) to permit the use of domain name registration data for any lawful purpose, it was not permitted to allow the use of that data to enable the transmission of mass, unsolicited, commercial email (spam). Register.com worked with business partners to provide web hosting and other Internet services to its domain name registration customers. Verio, a provider of web hosting and other Internet services, used a software robot to collect information about recently registered domain names and then contacted the individuals who had registered them to offer them various Internet services. The manner in which Verio contacted Register.com customers was calculated to cause Register.com customers to believe Verio was, at a minimum, a business partner of Register.com when in fact it was a competitor. After Register.com learned, through complaints from its customers and business partners, that Verio was soliciting its customers in this manner, it demanded that Verio stop making such solicitations. When Verio would not agree to cease the solicitations, Register.com sought an injunction to stop Verio, pleading trespass to chattels, breach of contract, unfair competition, and unauthorized access under the Computer Fraud and Abuse Act. The court issued a preliminary injunction based on a finding that Register.com was likely to prevail on both the trespass to chattels and computer fraud claims.

The holding in Register.com has been controversial because the court did not require a showing that Verio’s unauthorized access was actually interfering with the functioning of Register.com’s equipment. Instead, the court accepted Register.com’s argument that if Verio was allowed to continue this type of unauthorized access, the floodgates would open and there would be no end to the other companies using the same technique to harvest data from Register.com’s system, at which point the functioning of its equipment would be impaired.

If the court had used the constructive consent approach, the terms of the RAA, which required Register.com to provide public access to its data except under two limited circumstances, might have provided

Crafting a License

evidence of what access Register.com could be deemed to have consented. Instead, because Verio had no rights as a third-party beneficiary under the RAA, the court rejected evidence that Register.com’s terms of use of its data were more restrictive than the RAA permitted. However, subsequent to the Register.com lawsuit, ICANN revised the terms of the RAA, requiring domain name registrars not to give access to data for “transmission by e-mail, telephone, or facsimile of mass, unsolicited, commercial advertising or solicitations to entities other than the data recipient’s own existing customers.” This revised RAA closely resembles the Register.com terms in effect at the time of Verio’s unauthorized access, indicating that Register.com’s terms of use were not unreasonable even if they did not conform to its obligations under the version of the RAA in effect at the time of the litigation. If the later standard of acceptable use of Register.com data was used as the standard by which constructive consent should be measured, then Verio’s conduct would not have been privileged after all.

B. Ticketmaster Corp. v. Tickets.com, Inc.

Ticketmaster Corp. v. Tickets.com, Inc. considered the application of trespass doctrine to the problem of one commercial site linking to another without permission. Ticketmaster is the largest company selling tickets to sporting and other entertainment events, and has both online and bricks-and-mortar operations; Tickets.com is one of its competitors, operating primarily online. Tickets.com provided visitors to its own web site information about events for which Ticketmaster was the exclusive sales agent by providing “deep links” into Ticketmaster’s web site. These deep links permitted Tickets.com’s visitors to avoid Ticketmaster’s home page and directly access information about a particular event located deep within Ticketmaster’s web site. In 2000, Ticketmaster unsuccessfully sought a preliminary injunction to stop Tickets.com from providing deep links into its site, claiming breach of

54. Id. at 248 (stating that the RAA expressly provided that no third party beneficiaries would be created by its terms).
55. Internet Corporation for Assigned Names and Numbers Registrar Accreditation Agreement (ICAAN RAA) § 3.3.5 (May 17, 2001), available at http://www.icann.org/registrars/ra-agreement-17may01.htm.
57. Id.
contract, copyright infringement, trespass to chattels, false advertising, and trademark infringement.58

In 2003, the same court granted summary judgment to Tickets.com on Ticketmaster’s trespass to chattels and copyright claims, but denied summary judgment on the breach of contract claims.59 Tickets.com sent a software robot onto the Ticketmaster web site to copy information about events, then discarded information such as logos, advertisements, and formatting, keeping only the factual information describing the events. The court rejected the idea that any unauthorized access by a software robot could give rise to liability for trespass when no impairment of the function of the equipment had been shown.

The notion of constructive consent would point toward the same result here because hypertext is one of the defining characteristics of the World Wide Web. The court found that Ticketmaster had not shown that it had suffered any damages, such as loss of advertising revenues, as a result of the deep links created by Tickets.com. While the holding in the case was based on the inability of Ticketmaster to make out any trespass claim at all based on failure to prove the element of damages, the lack of damages also supports a finding that the access by Tickets.com through deep linking was the functional equivalent of reasonable access by an individual user. Ticketmaster wanted anyone interested in events for which it was the exclusive agent to access its site from its home page, but when the original dispute arose in 1999, it had not implemented any access controls that would have required individual visitors to follow such a route.

C. Intel Corp. v. Hamidi

In Intel Corp. v. Hamidi, a disgruntled former employee used the Internet to criticize his erstwhile employer by sending emails to current Intel employees and by building a web site to disseminate his anti-Intel opinions on the World Wide Web.60 After someone provided him with an electronic copy of Intel’s employee directory, Hamidi sent emails on

59. Ticketmaster Corp., 2003 U.S. Dist. LEXIS 6483, at *2 (finding immunity from liability). In 2000, the trial court found that Ticketmaster was unlikely to prevail on its breach of contract claim, but after additional discovery turned up new evidence that Tickets.com may have entered into a contract with Ticketmaster over the Internet the court in 2003 refused to enter summary judgment in favor of Tickets.com.
60. 71 P.3d 296 (Cal. 2003).
Crafting a License

six occasions to anywhere from 8,000 to 35,000 employees, or a total of between 48,000 and 210,000 emails in all.\textsuperscript{61} Hamidi’s messages promised to remove recipients from the mailing list upon request, and he apparently complied with all such requests that he received.\textsuperscript{62} Intel, however, was unwilling to wait for its employees to make such requests or simply to delete Hamidi’s emails from their inboxes, and so demanded that Hamidi stop sending critical emails to its current employees.\textsuperscript{63} At trial and on appeal,\textsuperscript{64} Intel’s request for an injunction was granted based on a trespass to chattels theory, notwithstanding the lack of a showing of any significant impairment of Intel’s system’s functioning. The California Supreme Court took a different view of trespass doctrine, and held that because there was no significant impairment of Intel’s system’s functioning, and because there was no indication that a flood of other former disgruntled employees were waiting to deluge it with emails, Hamidi was immune from liability under trespass doctrine.\textsuperscript{65}

In determining whether Intel should be deemed to have consented to Hamidi’s use of its equipment to send emails to its current employees, notwithstanding its vociferous objections, a crucial factor would seem to be the non-commercial character of Hamidi’s communication. Although the California Supreme Court did not reach the issue of whether enforcement of a state law that interfered with Hamidi’s exercise of his free speech rights would constitute impermissible state action under the First Amendment, unlike the \textit{Register.com} and \textit{Ticketmaster} cases, Hamidi was not a competitor of Intel, and his behavior did not raise unfair competition issues. On the other hand, if constructive consent creates a privilege for reasonable individual access of Internet facilities, it is unclear whether sending 48,000 to 210,000 emails constitutes reasonable individual use. A final factor suggesting that Intel should be deemed to have constructively consented to Hamidi’s sending emails to its current employees is its failure to implement more effective and restrictive access controls. By granting its employees relatively free access to the Internet, Intel arguably entered into a public arena within

\textsuperscript{61} Id.
\textsuperscript{62} Id.
\textsuperscript{63} Id.
\textsuperscript{64} Intel Corp. v. Hamidi, 114 Cal. Rptr. 2d 244 (Cal. Ct. App. 2001), rev’d, 71 P.3d 296 (Cal. 2003).
\textsuperscript{65} Hamidi, 71 P.3d at 311.
which it could be expected to tolerate the criticism of ex-employees such as Hamidi, at least in the absence of any showing that Hamidi’s communications were impairing the function of its systems.

IV. CONCLUSION

The public character of the Internet can be protected by assigning a legal significance to decisions by commercial Internet users about how to make use of Internet technologies. Individual users of the Internet should have a license to access what is posted on the Internet that cannot be negated by arbitrary assertions of rights over information rooted in ownership of tangible computer equipment. Because the free flow of information has been a hallmark both of civil society and the Internet, that association between civil liberties and the institutional character of the Internet should be preserved notwithstanding the growing commercialization of Internet resources. Granting individual users a privilege to access information on the Internet in a reasonable manner would preserve the basic character of that association while still recognizing that the Internet is now sustained by private investment in computer equipment. Access by anyone other than individuals in a manner that approximated individual access would likewise be covered by an implied license created by constructive consent, but consent to access that differs in quality or quantity from that associated with individual users could be withdrawn. This license to access information would be in effect a privilege implied by law that limits the property rights of the owners of the equipment. Using the common law to articulate the scope of constructive consent to access by Internet users, and the privilege it creates, would help to clarify the social significance of the Internet itself and establish viable standards to safeguard its open, participatory character.
ONLINE COURT RECORDS: BALANCING JUDICIAL ACCOUNTABILITY AND PRIVACY IN AN AGE OF ELECTRONIC INFORMATION

Peter A. Winn*

I. INTRODUCTION

Francis Bacon said that knowledge is power.¹ He could just as easily have identified power with the control of information. As Bacon—the lawyer, judge, and Lord Chancellor—well knew, courts were and still are the pre-eminent information systems— institutions that process information and translate it into the exercise of power; in the case of courts, by rendering judgments. Later Enlightenment thinkers such as Cesare Beccaria well understood the connection between judicial information and power, and argued forcefully that all trials should be public.² Publicity was viewed as a check against the misuse of judicial power, tending to limit unfair (or at least unpopular) prosecutions by the rulers of a society, as well as increasing public respect for the legal system.³ Of course, while the legal system has inherited from the Enlightenment a presumption of openness, that presumption has always been limited when unfair publicity, itself, threatens to become an

* Senior Fellow, University of Washington School of Law; Lecturer, Graduate Studies Department, University of Melbourne School of Law; Assistant U.S. Attorney, U.S. Department of Justice. The views expressed in this Article are the author’s personal views and should not be interpreted as reflecting the position of the U.S. Department of Justice. The author wishes to express his gratitude to the following persons for their comments: Judge Donald Horowitz, Judge Dennis Michael Lynn, Dan Solove, Robert Ellis Smith, Gregory Silverman, and Katie Simon.

¹. See JOHN BARTLETT, FAMILIAR QUOTATIONS 111 (Christopher Morley & Louella D. Everett eds., 11th ed., Little, Brown & Co. 1940) (1882) (attributing quote “Nam et ipsa scientia potestas est” (knowledge is power) to Francis Bacon in his work Meditationes Sacrae, De Haeresibus). The use of the term potestas by Bacon was no accident. Potestas is used to denote legal power or control, and Bacon, in his professional life, served in many different judicial capacities, including that of Lord Chancellor.


³. “Let the verdicts and proofs of guilt be made public, so that opinion, which is, perhaps, the sole cement of society, may serve to restrain power and passions; so that the people may say, we are not slaves, and we are protected—a sentiment which inspires courage and which is the equivalent of a tribute to a sovereign who knows his own true interests.” Id. at 22.
instrument of oppression. Experience teaches us that at times, open judicial proceedings can ensure, rather than prevent, the abuse of judicial power, can create unacceptable risks of a miscarriage of justice, and can cause unnecessary harm to the safety and privacy of individuals.

This Article examines the traditional balance courts have reached between the disclosure of information generated by the judicial process and the need at times to limit the disclosure of that information. The Article then examines how this traditional balance is upset when judicial information is placed online. The Article argues that as courts adapt to a world of electronic information, new rules and practices must be established to maintain the policies underlying the traditional balance. While there must continue to be a presumption of openness, courts must limit the disclosure of judicial information when it threatens the effective administration of justice and when necessary in order to protect the safety and privacy of individuals participating in the judicial process.

A. Access to Information in Criminal Proceedings

The presumption of openness of judicial proceedings is embodied in the Sixth Amendment to the U.S. Constitution, which guarantees the accused in every criminal case the right to a public trial. In the words of Justice Hugo Black, the Sixth Amendment is “a safeguard against any attempt to employ our courts as instruments of persecution. The knowledge that every criminal trial is subject to contemporaneous review in the forum of public opinion is an effective restraint on possible abuse of judicial power.”

At the same time, the presumption of openness is limited when it interferes with the fair and impartial administration of justice, or threatens the safety or the reasonable expectation of privacy of the participants in the judicial process. Thus, there is no general public right of access to federal grand jury proceedings, which under Federal Rule

---

4. U.S. CONST. amend. VI (“In all criminal prosecutions, the accused shall enjoy the right to a speedy and public trial, by an impartial jury of the State and district wherein the crime shall have been committed, which district shall have been previously ascertained by law, and to be informed of the nature and cause of the accusation; to be confronted with the witnesses against him; to have compulsory process for obtaining witnesses in his favor, and to have the Assistance of Counsel for his defense.”).


Online Court Records

of Criminal Procedure 6(e) are conducted in secrecy. This rule protects the reputations of individuals if they are innocent, and limits their opportunities for obstruction of justice if they are guilty. Applications for search warrants and for electronic surveillance nearly always take place in secret. Finally, judges have the discretion to close the courtroom to public spectators when necessary to achieve a fair trial—for instance, when faced with witness intimidation or other disruption to the judicial proceedings. The Sixth Amendment confers no special benefits on the press, and the right to a public judicial proceeding may be waived by a defendant—one who, for instance, may wish to limit the risk that he may be harmed if his counterparts learn of an agreement to cooperate with the government’s investigation. Many criminal court records are prepared and maintained entirely in secret. For example, there is no right of public access to pre-sentence reports that are prepared by the federal probation office and contain extremely sensitive information about criminal defendants’ health, mental condition, family history, and finances.

In addition to the Sixth Amendment right of a criminal defendant to a public trial, the U.S. Supreme Court has also recognized a limited right of access to criminal judicial proceedings under the First Amendment of the Constitution. In Richmond Newspapers, Inc. v. Virginia, the Court held that “[i]n guaranteeing freedoms such as those of speech and press, the First Amendment can be read as protecting the right of everyone to attend trials so as to give meaning to those explicit guarantees.”

8. See, e.g., Sells Eng’g, Inc., 463 U.S. at 424 (“Grand jury secrecy, then, is ‘as important for the protection of the innocent as for the pursuit of the guilty.’” (quoting United States v. Johnson, 319 U.S. 503, 513 (1943))).
11. See United States v. Akers, 542 F.2d 770, 772 (9th Cir. 1976).
16. Id. at 575; see also Press-Enter. Co. v. Superior Court, 478 U.S. 1, 11–12 (1986).
Nevertheless, the right of access to the courts is not absolute. Courts continue to have the duty to balance the presumption in favor of public access against other interests that may justify restricting access. These include the possibility of prejudicial pretrial publicity; the danger of impairing law enforcement or judicial efficiency; and the protection of the legitimate privacy interests of litigants and other persons, such as witnesses, victims, and jurors.17

B. Access to Information in Civil Proceedings

In the context of civil proceedings, it has been held that the Sixth Amendment does not support a general constitutional right of access.18 Nevertheless, for many of the same underlying reasons supporting public access to criminal proceedings, the U.S. Supreme Court has recognized a common law right “to inspect and copy public records and documents, including judicial records and documents.”19 In Nixon v. Warner Communications, Inc.,20 Justice Powell, writing for the Court, articulated the limits on this common law right of access: “It is uncontested, however, that the right to inspect and copy judicial records is not absolute. Every court has supervisory power over its own records and files, and access has been denied where court files might have become a vehicle for improper purposes.”21

In Nixon, the Court considered whether the press could obtain copies of tape recordings of conversations between former President Nixon and various members of his staff that had been introduced into evidence in the trials of these staff members.22 Although these tape recordings had been played in public during the trial, until then the press had only been able to obtain transcripts of the tapes.23 The Court held that after the conclusion of the judicial proceedings, Nixon’s interest in privacy outweighed the common law right of the press to have copies of the

21. Id. at 598.
22. Id. at 589.
23. Id.
Online Court Records

tapes, particularly when the only purpose that could be cited for the release of the copies was their potential for commercial exploitation.  

As stated by the Court in *Nixon*, the underlying purpose of the common law right of access is the “citizen’s desire to keep a watchful eye on the workings of public agencies and in a newspaper publisher’s intention to publish information concerning the operation of government.” In another leading case, *In re Continental Illinois Securities Litigation*, the U.S. Court of Appeals for the Seventh Circuit stated that the purpose of the common law right of access is to allow the citizenry to “monitor the functioning of our courts, thereby insuring [sic] quality, honesty, and respect for our legal system.” This rationale, of course, comes directly from Enlightenment thinkers such as Beccaria.

C. Balancing Access and Privacy

Understanding the rationale for public access to judicial proceedings, however, also reveals the limitations of that rationale. These limitations become evident when courts must balance the presumption of the openness of judicial proceedings against the need to keep certain types of information confidential—in particular, sensitive personal information. Courts tend to favor public access when the underlying purpose of public access is to ensure the integrity of the judicial process. On the other hand, courts tend to protect personal information when the purpose of access is not related to facilitating public scrutiny of the judicial process, but to exploiting information in judicial records for commercial or other purposes unrelated to public oversight of the judicial system.

The balance worked out by the courts bears a close analogy to the concept of fair information practices as applied in the context of the federal Privacy Act and the Freedom of Information Act (FOIA).

---

24. *Id.* at 602.
25. *Id.* at 598 (citations omitted).
26. 732 F.2d 1302 (7th Cir. 1984).
27. *Id.* at 1308.
28. See supra note 2 and accompanying text.
information for one purpose is not ordinarily permitted to use that information for a different unrelated purpose. Public access to sensitive personal information about individuals is generally prohibited unless it serves the purpose of ensuring accountability in government. In a very similar manner, the cases involving access to court records evidence a careful effort to balance the presumption of public access and the privacy rights of individuals. In framing this balance, courts are sensitive to protect not only the personal privacy of litigants, but also the harm that can come to others, such as witnesses, victims, jurors, and other third parties, who may have no control over the information so disclosed.

The sensitivity shown by court decisions in balancing access and privacy in the context of judicial records reflects the large body of judicial experience addressing questions of privacy in other contexts. Courts are accustomed to balancing the social benefits from the disclosure of personal information against the risk of harm that such disclosure may cause the individuals who are so identified. For example, in Nixon v. Administrator of General Services, the U.S. Supreme Court determined that President Nixon had a constitutional privacy interest in records of his private communications with his family, but not in records of his official duties.

Courts have also shown particular sensitivity in protecting personal health information from disclosure. In Whalen v. Roe, the U.S. Supreme Court recognized that because of its great sensitivity, personal health information is protected under a constitutional right to information privacy. However, in Whalen, the Court permitted the collection of this personal health information for purposes of public health and safety when there were strong and effective assurances that the information so collected would be kept confidential. In these cases,
Online Court Records

courts have excelled in demonstrating a great capacity for careful and nuanced balance.41

In the leading case of *Westinghouse Electric Corp. v. United States*,42 the U.S. Court of Appeals for the Third Circuit set out five factors that must be considered in determining the appropriate constitutional balance between personal privacy and a governmental interest in disclosure of health records: (1) the type of health record requested and the type of health information it contains, (2) the potential for harm in any subsequent non-consensual disclosure of the information, (3) the injury from disclosure to the relationship in which the record was generated, (4) the adequacy of safeguards to prevent unauthorized disclosure, and (5) the degree of need for access.43 Relying on the *Westinghouse* test, some courts have found a constitutional tort under 42 U.S.C. § 198344 for the improper disclosure of health information by state government officials.45 On the other hand, in *Doe v. Southeastern Pennsylvania Transportation Authority*,46 the same appellate court found that the disclosure of medical records to the administrator of a health benefit plan was not actionable.47

The pragmatic reasons supporting the need for public access (for example, the need to assure credibility and accountability of the judicial system) are typically balanced against the pragmatic reasons supporting the need to restrict public access (for example, protecting the rights of litigants to a fair trial, protecting the rights of individuals to privacy, and protecting individuals from harm caused by misuse of their personal information). While courts are vigilant in protecting the public right of access when it is consistent with ensuring the credibility of the judicial

42. 638 F.2d 570 (3d Cir. 1980).
43. *Id.* at 578.
45. *Doe v. Borough of Barrington*, 729 F. Supp. 376, 378, 382 (D.N.J. 1990) (finding a violation of the plaintiff’s constitutional right to privacy when a police officer who had arrested the plaintiff disclosed to the plaintiff’s neighbor the fact that the plaintiff had HIV); *Woods v. White*, 689 F. Supp. 874, 876 (W.D. Wis. 1988) (holding that the constitutional right of privacy extended to the fact that a prison inmate had tested positive for HIV where allegedly disclosed by prison medical personnel to non-medical staff and other inmates), *aff’d without opinion*, 899 F.2d 17 (7th Cir. 1990).
46. 72 F.3d 1133 (3d Cir. 1995).
47. *Id.* at 1143.
system, they are also quick to protect individuals from the exploitation of their personal information when it bears little relationship to ensuring the integrity of the judicial process. This common law and constitutional balance, carefully worked out on a case-by-case basis over the course of many years, represents the finest form of judicial lawmaking. While a system that relies on the discretion of judges sometimes runs the risk of occasional inconsistent decisions, by and large, courts have shown that they are capable of exercising their discretion to carefully weigh competing interests, and their decisions show great nuance, factual subtlety, and legal imagination.

II. ELECTRONIC JUDICIAL RECORDS: HOW TO MAINTAIN BALANCE

A. The Move from Paper Records to Electronic Records

When courts move from the world of paper judicial records to the world of electronic judicial records, they must confront the same issues underlying the information revolution throughout our society. The revolution in the use of electronic information has seen the capacities of hardware, software, and communications networks continually increase and their costs continually decrease. This has permitted information to be used in ways that were previously impractical. In the law, the conversion from paper to electronic judicial records has provided courts and attorneys the opportunity to obtain substantial benefits in the operation of the legal system. The conversion also offers the public the opportunity to better understand and appreciate the judicial process.

However, as our legal system undergoes the transformation to a system of electronic judicial records—with all its substantial benefits—it

48. See supra notes 25–29 and accompanying text.
49. See supra notes 21–24, 30 and accompanying text.
50. Professor Solove, for instance, takes the position that the current system does not provide enough protection for individual privacy. See Daniel J. Solove, Access and Aggregation: Public Records, Privacy and the Constitution, 86 MINN. L. REV. 1137, 1154–59 (2002).
51. The federal electronic case filing system offers the following benefits: 24-hour access to case files over the Internet, ability to file pleadings electronically with the court, automatic e-mail notice of case activity, ability to download and print up-to-date documents directly from the court system, no waiting in line, expanded search and reporting capacities, the elimination of the cost of expensive courier services, and an overall reduction in the physical storage space needs of the courts. See Federal Public Access to Court Electronic Records ("PACER") system website, at http://pacer.psc.uscourts.gov/cmecf (last visited Jan. 5, 2004).
Online Court Records

is critical to ask how the advantages of public access are to be balanced against the other competing policies that have served to limit access in the past, such as maintaining the integrity of the judicial system and protecting individuals from invasion of their privacy and misuse of their personal information. It is temptingly easy to assume that if one applies the same set of rules to electronic judicial records that was applied in the past to paper records, it will result in the same balance between the various competing policies. Unfortunately, this is not the case. The assumption of parity represents a serious misunderstanding of the differences between paper records and electronic records. When the same rules that have been worked out for the world of paper records are applied to electronic records, the result does not preserve the balance worked out between the competing policies in the world of paper records, but dramatically alters that balance. It shifts the balance away from individual privacy, producing little if any benefit on the side of judicial accountability.

In this context, to assert that electronic judicial records should be placed under the same rules as paper records is nothing more than to advocate for the free flow of information at the expense of the many other competing values. This position appears to be driven more by an abstract philosophy about the importance of the free flow of information than a study of the actual historical experience of courts. That experience forces us to recognize that the flow of information can have very serious and concrete consequences—not all of them beneficial to the effective administration of justice or the protection of the dignity and security of individuals. For instance, the unrestricted shift to electronic court records permits the type of commercial exploitation of judicial records that courts have traditionally eschewed. Such commercial exploitation of judicial records harms the privacy of litigants with virtually no corresponding benefit to the administration or the accountability of the justice system. Furthermore, if the shift from paper to electronic court records takes place without appropriate safeguards, we will celebrate the abstract value of the free flow of judicial information at the cost of the privacy and security of litigants and other participants in the judicial system. They will lose not only their interest in privacy—their identities will be subject to potential misuse by thieves, and their children may be exposed to sexual predators. Instead of increasing social respect for the judicial system, unrestricted access to court records will undermine the respect and confidence the courts in this country have traditionally enjoyed.
B. The “Practical Obscurity” of Paper Records also Protects Privacy

In the past, paper-based records, while technically public, continued to retain a high degree of “practical obscurity.” This was true because the retrieval of paper records involved costs that do not attend the retrieval of electronic records. In the world of paper judicial records, personal information could be open to the “public” in the sense that it could be accessed by any member of the general public, but the costs of retrieval limited access as a practical matter. Only those with a relatively strong interest in the information would take time out of their day, wait in line at the clerk’s office, fill out the necessary forms, and pay the necessary copy charges. Once judicial records go online, however, computerized compilers can search, aggregate, and combine the information with information from many other public filings to create a profile of a specific individual in a matter of minutes, at minimal cost. Information in many different locations can be combined and aggregated in ways that previously were impossible, permitting entirely new uses of the information that could never have been intended before. These search, aggregation, and bulk dissemination capabilities can handle major magnitudes of information, take little time, and have minimal cost. While there may be social benefits from this increased access to personal information, electronic access will also mean that court records can be commercialized in ways they have never been before.

Paper records also exist in time and space differently from electronic records. Paper records—like human beings—are organic. As such, they experience a natural progression of decay and change. Over time, paper-based information accumulates and grows old and must be cleared away to make room for the new. The “practical obscurity” of old records generates an expectation of privacy that has been recognized as legitimate by common law courts. The concept of practical obscurity developed by federal courts in the context of FOIA cases also recognizes that the passage of time may actually increase the privacy interest at stake when disclosure would revive information that was once public.

53. See Solove, supra note 50, at 1149–52.
54. See, e.g., Briscoe v. Reader’s Digest Ass’n, Inc., 483 P.2d 34, 36, 44 (Cal. 1971) (holding that a truthful publication of an eleven-year old criminal conviction constitutes a valid cause of action for invasion of privacy); Melvin v. Reid, 297 P. 91, 91, 93 (Cal. Ct. App. 1931) (holding that a movie truthfully depicting plaintiff’s previous life as a prostitute many years earlier constitutes a valid cause of action for invasion of right to pursue happiness).
knowledge but has long since faded from memory. On the other hand, electronic records are inorganic; they do not grow old, get moved to warehouses, or eventually get destroyed. They continue to exist, potentially forever. In an age of electronic information, a serious question arises as to whether a rehabilitated criminal will be allowed to put his past behind him, whether a former prostitute who was acquitted of a murder charge will ever be allowed to forget it, or whether a victim of a sexual assault will be allowed to heal her wounds and not be victimized once again by reminder and new public disclosure many years later.

C. The Dark Side of Online Access

In this context, it may be helpful to examine other examples of online information. Putting consumer credit information online has permitted credit-reporting agencies to aggregate huge amounts of personal financial information. While this has created unprecedented access to consumer credit in the United States, it also has spawned a new type of crime—identity theft. Identity theft is now reaching epidemic proportions and has left millions of innocent victims little, if any, means of redress. In health care, the electronic revolution has created new opportunities for advances in medicine, but it also has begun to

55. See, e.g., Reporters Comm. for Freedom of the Press, 489 U.S. at 767 (“O[ur cases have also recognized the privacy interest inherent in the non-disclosure of certain information even when the information may have been at one time public.”); Rose v. Dep’t of the Air Force, 495 F.2d 261, 267 (2d Cir. 1974) (“[A] person’s privacy may be as effectively infringed by reviving dormant memories as by imparting new information.”), aff’d, 425 U.S. 352 (1976).

56. See Briscoe, 483 P.2d at 36.

57. See Melvin, 297 P. at 91.


undermine the relationship of trust between physician and patient. 60
Congress has attempted to address some of the social problems created
by electronic financial and health care records through the enactment of
privacy legislation. 61

In Section 205 of the E-Government Act of 2002, Congress directed
the federal court system to implement public access to the Internet by
2004 and also directed the Judicial Conference to promulgate rules to
address concerns about the privacy and security of personal information
in light of the best practices of the federal and state courts. 62 The E-
Government Act places considerable discretion in the hands of the
courts, and indicates a congressional deference to the courts to be
responsible for the management and oversight of their own records. 63
However, the language in the E-Government Act clearly shows that
Congress believed that, as always, courts continue to have a
responsibility to balance the benefits of public access against the concern
for the security and privacy of individuals.

It is also important to remember in this context that courts have a
responsibility not only to protect the litigants—most of whom at least
are represented by counsel (although an increasing number of litigants
are acting pro se)—but also a responsibility to witnesses, victims, jurors,
and other third parties who enter the legal system without the due
process protections of the Fourteenth Amendment. 64 Courts must ensure
that those who encounter the legal system—voluntarily as well as
involuntarily—do not face exploitation of their personal information by
commercial information brokers or become the victims of cyber-
criminals or electronic peeping toms.

D. Legal Protections for the Privacy Value in Practical Obscurity

The information age has transformed the world so quickly that there
has been little time to develop case law that adjusts the balance between
public access and the protection of personal information with the

60. See Peter A. Winn, Confidentiality in Cyberspace: The HIPAA Privacy Rules and the
61. Right to Financial Privacy Act, 12 U.S.C. §§ 3401–3422 (2000); Fair Credit Reporting Act,
63. Id.
64. U.S. CONST. amend. XIV.
Online Court Records

differences between electronic and paper records in mind. However, from the limited number of cases addressing the question of computerized information, there is a growing recognition by courts that a very different balance must be applied. In Whalen, the fact that the information was compiled in a mainframe computer database was of tremendous significance to the Court.65 Although the Court found that the system of protections established by the government passed constitutional muster,66 Justice Brennan wrote a concurrence that was prescient as to the future problems to come:

What is more troubling about this scheme, however, is the central computer storage of the data thus collected. Obviously, as the State argues, collection and storage of data by the State that is in itself legitimate is not rendered unconstitutional simply because new technology makes the State’s operations more efficient. However, as the example of the Fourth Amendment shows, the Constitution puts limits not only on the type of information the State may gather, but also on the means it may use to gather it. The central storage and easy accessibility of computerized data vastly increases the potential for abuse of that information, and I am not prepared to say that future developments will not demonstrate the necessity of some curb on such technology.67

The case that is perhaps most applicable to the problem of electronic judicial records is United States Department of Justice v. Reporters Committee for Freedom of the Press.68 In this FOIA case, Justice Stevens cited Whalen in rejecting the press’s argument that because a person’s FBI computerized rap sheet was merely a summary of public records of arrests and convictions from local state and county courthouses around the country, the FBI computerized rap sheet itself should be a public document accessible under the FOIA.69 He observed: “Plainly there is a vast difference between the public records that might be found after a diligent search of courthouse files, county archives, and local police stations throughout the country and a computerized summary located in a single clearinghouse of information.”70

66. Id. at 600.
67. Id. at 606–07 (Brennan, J., concurring).
69. Id. at 762–63.
70. Id. at 764.
Strictly speaking, of course, the test established by Justice Stevens in *Reporters Committee for Freedom of the Press* applies only in a FOIA context to the FBI’s record-keeping system (although it involved court records). However, the underlying problems in that case—the problems of balancing public access to information against concerns for individual privacy—are exactly the same. Thus, the balance reached by the Supreme Court in that case represents an extremely important precedent. The Court defined the public’s interest as shedding “light on the conduct of any Government agency or official” rather than acquiring information about a particular private citizen. 71 The Court also noted “the fact that ‘an event is not wholly ‘private’ does not mean than an individual has no interest in limiting disclosure or dissemination of the information.’” 72 As in the case of the rap sheets addressed in *Reporters Committee for Freedom of the Press*, before the advent of electronic case files, the right to “inspect and copy” court files depended on physical presence at the courthouse. The inherent difficulty of obtaining and distributing paper case files largely insulated litigants and third parties from the harm that could result from massive or unnecessary exposure, dissemination, or misuse of information provided in connection with a legal proceeding. When the U.S. Supreme Court has focused on the new world of electronic information, it recognized that it had to alter the balance formerly adopted between the competing policies of granting and limiting access to judicial records.

E. The Way Forward

If the message of *Reporters Committee for Freedom of the Press* is followed by courts about to adopt a system of electronic judicial record-keeping, it would appear that there must be a new balancing that takes place. Courts must recognize that their case files often contain private or sensitive personal information—such as medical and health records, employment records, detailed financial information, tax returns, Social Security numbers, intimate family information, intimate victim information, and other personal and identifying information. In the world of paper judicial records, the inconvenience of access provides considerable practical protection for the concerned individuals.

71. *Id.* at 773.

72. *Id.* at 770 (quoting Rehnquist, *Is an Expanded Right of Privacy Consistent with Fair and Effective Law Enforcement?*, Nelson Timothy Stephens Lectures, University of Kansas Law School, pt. 1, p. 13 (Sept. 26–27, 1974)).
Online Court Records

However, when courts permit these case files to become electronic and connected to the Internet without proper safeguards, they will make all this personal information available easily and almost instantly for downloading, storage, searching, data compilation, aggregation, and massive dissemination for purposes that were never intended by either litigants, witnesses, victims, jurors, or others involved with or connected to a court proceeding.

F. Maintaining the Balance in the Movement to Online Court Records

Reaching a balance between judicial accountability and the protection of privacy was not an easy task in the days of paper-based judicial records. The cases reflect a continual struggle between the goals of ensuring open access to court records and the competing goals of ensuring the effective working of the judicial system and respecting the privacy of the participants. The decisions are fact specific and highly contextual, but they all reflect sensitivity by the courts to the importance of balancing the various competing interests at stake. In the world of electronic information, the conflict between principles of open access and privacy becomes much more sharp and difficult for courts to resolve. There is no easy solution. Ideally, courts should be free to engage in the same common law decision-making process that courts used to resolve such conflicts in the age of paper-based judicial records.

Unfortunately, the most important decisions will not take place as a matter of case-by-case decision making, but will be the decisions that courts make as an administrative matter when they select the computerized record-keeping system they will use and when they establish the rules governing the use of that system. The balance must be worked out in the way administrative agencies make decisions—presumably after notice, public participation, comment, and thoughtful reflection on the costs and benefits of the various alternatives. A thoughtful set of guidelines for public access to court records has been produced by the National Center for State Courts and the Justice Management Institute. In the same vein, the Canadian Judicial Council

has released a very thoughtful discussion paper on open courts, electronic access to court records, and privacy.74

G. The Experience of the Federal Courts

Perhaps the most significant development in the context of this debate is the Report of the Judicial Conference Committee on Court Administration and Case Management on Privacy and Public Access to Electronic Case Files (Report).75 Before the Judicial Conference Committee on Court Administration and Case Management (Committee) issued the Report, a study of the problem was prepared by the staff of the Administrative Office of the United States Courts.76 The staff white paper described two general approaches to the problem.77 One approach was to treat electronic judicial records as governed by exactly the same rules as paper records—what the white paper calls the “public is public” approach.78 The second approach advocated treating electronic and paper files differently in order to respect the practical obscurity of paper case files, urging that the rules regulating electronic court records reflect the fact that unrestricted online access to court records would undoubtedly, as a practical matter, compromise privacy, as well as increase the risk of personal harm to litigants and third parties whose private information appeared in case files.79 The white paper suggested that different levels of privileges could be created to govern electronic access to court records.80 Under this approach, judges and court staff would generally have broad, although not unlimited, remote access to all electronic case files, as would other key participants in the judicial process, such as the

77. Id. at 7.
78. Id.
79. Id. at 7.
80. Id. at 10.
Online Court Records

U.S. Attorney, the U.S. Trustee, and bankruptcy case trustees. Litigants and their attorneys would have unrestricted access to the files relevant to their own cases. The general public would have remote access to a subset of the full case file, including, in most cases, pleadings, briefs, orders, and opinions. Under this approach, the entire electronic case file could still be viewed at the clerk’s office, just as the paper file is available now for inspection, but would not generally be made available on the Internet.

Unfortunately, at least with respect to civil cases and bankruptcy cases, few, if any, of the suggestions contained in the staff white paper were ultimately adopted in the Report. Instead, the Committee adopted the “public is public” approach to the problem, rejecting the view that courts have a responsibility to adopt rules governing the use of their computer systems to try to recreate in cyberspace the practical balance that existed in the world of paper judicial records. In supporting this decision, the Committee took the position that attempting to recreate the “practical obscurity” of the brick and mortar world was simply too complicated an exercise for the courts to undertake. The Report does appear to recognize a limited responsibility on the part of the courts to adopt rules in order to limit the foreseeable harms of identity theft and online stalking. The Report recommends that certain “personal data identifiers,” such as Social Security numbers, dates of birth, financial account numbers, and names of minor children, be partially redacted by the litigants.

With respect to the problem of protecting individual privacy, the Report places the burden on parties and counsel to anticipate these questions, and advises them to use motions to seal and for protective orders on a case-by-case basis. Although it is reasonable to hold the

81. Id.
82. Id.
83. Id.
84. Id.
85. See REPORT, supra note 75, at 3–7.
86. See id. at 4.
87. Id.
88. See id. at 6.
89. Id. at 3–5. Social Security case files were excluded entirely from online access, presumably because of the large amount of personal health information contained in such files. Id. at 5.
90. Id. at 3. The Committee’s Report recognizes that the public access requirement of the Bankruptcy Code, 11 U.S.C. § 107 (2002), appears to preclude bankruptcy courts from sealing any
parties and their attorneys primarily responsible to protect their own privacy, the Report could have and should have done more in this respect. For instance, its guidelines could have included more explicit warnings to attorneys to exercise caution when filing any personal identifying number, such as a driver’s license number, medical records, treatment and diagnosis, employment history, individual financial records, information pertaining to children, or proprietary or trade secret information. The most significant weakness in the Report is that it leaves unanswered the question of how the system will protect the privacy of pro se litigants or third parties who are not litigants or have not voluntarily chosen to enter the justice system—foremost among these are jurors, witnesses, victims of crimes, and their family members.

The Report recommends that criminal court records not be placed online, for the present, finding that any benefits of remote electronic access to criminal files would be outweighed by the safety and law enforcement risks such access would create. The Report expressed the concern that allowing defendants and others easy access to information regarding the cooperation and other activities of co-defendants would increase the risk that the information would be used to intimidate, harass, and possibly harm victims, defendants, and their families. In addition, the Report noted that merely sealing such documents would not adequately address the problems of online access, since the fact that a

court records to protect the privacy of individuals. REPORT, supra note 75, at 6. The Committee responds to this concern by suggesting that Congress amend 11 U.S.C. § 107 to address this problem. Id. In the meantime, it would presumably be possible for a court to order that certain records be filed publicly in the clerk’s office, but kept off-line. In addition, it would appear to be possible, as some bankruptcy courts have done, to adopt local rules requiring certain types of information—such as Social Security numbers, numbers of financial accounts, and names of minor children—not be placed in the court file at all, while still providing this information to key participants in the system. Id.


92. The problem of protecting innocent third parties from harm is particularly pressing in bankruptcy cases. For instance, in large Chapter 11 cases, “First Day” orders typically provide for the payment of back wages of employees of the debtor. These orders typically contain large amounts of sensitive personal information, including salaries, health benefit information, and Social Security numbers. Innocent third parties, usually family members, are also usually implicated in individual bankruptcy cases.

93. REPORT, supra note 75, at 5.

94. Id.
document is sealed signals probable defendant cooperation and covert law enforcement initiatives.  

Of course, as federal courts have converted to electronic filing systems, criminal records are filed electronically as well. However, for criminal cases, access to electronic criminal records is controlled through a system of privileges. Judges and law clerks have full access to criminal court records, as do defense attorneys and prosecutors in those cases in which they are counsel. While members of the general public do not have online access to the criminal court files, copies of indictments and other documents that were publicly available in the past continue to be publicly available—by request in person at the clerk’s office. There is no technological reason why, on a case-by-case basis, a third party—such as a newspaper reporter or other interested member of the public—could not request similar electronic access to the case file. However, this access would be permitted only after notice to the other parties in the case who could be heard if electronic public access might involve any special risks.

From the limited experience in criminal cases, it appears that courts do have the capacity to implement an electronic system that uses different levels of privileges to recreate in cyberspace a system that protects the same values that were protected by the practical obscurity of paper records. Implementing a system of privileges does not appear to present any serious technological problems—the architecture of the federal PACER system is the same for both civil and criminal judicial records. However, if this is a correct evaluation of the current technology used by clerks in federal courts, a somewhat disturbing conclusion follows. It would appear that the Committee, at least in the case of civil and bankruptcy cases, simply made a decision to ignore the legal decisions with respect to the existence of a right to privacy in the practical obscurity of judicial records—or to have made a determination that these cases represented the wrong public policy. If so, one would hope that the Committee could have expressed the underlying basis of its decision more clearly. As it is, the Report leaves an outside observer with the disquieting feeling that the Committee may be using technology not as a reason, but as an excuse.

Under the mandates of the E-Government Act, the federal judiciary will continue to review and monitor the process of online access. The

95. Id.
Report itself recognizes that the federal courts will ultimately have to revise their rules and procedures about electronic access to judicial records in the light of experience.97 In doing so, the Committee will have to consider the concrete experience of both federal and state courts as they experiment with different models of information protection. In the meantime, both federal and state courts are likely to see considerably more litigation regarding motions to seal records and for protective orders. In the context of this litigation, parties desiring to protect personal information but unable to meet the standard for a seal or a protective order, may consider requesting the intermediate relief that documents filed electronically in their civil cases simply be kept off-line, in the same manner as criminal case files in the federal court. Under this approach, the parties might continue to have electronic access to their court file, and the court file would continue to be available for public viewing at the clerk’s office, or be made available electronically on the filing of a notice of appearance by any interested person. The parties should have the right to request that their personal information—and the personal information of others—simply not be made available for unrestricted access on the Internet.

H. Tentative Suggestions

In ruling on motions to protect personal information from online invasions of privacy, as well as in the process of fashioning local rules governing online access, the Westinghouse test98 may prove to be a model of how courts may have to determine what type of information parties should be permitted to protect. Is it the type of information that society as a whole is prepared to recognize as involving a reasonable expectation of privacy? What is the potential for harm to the person in the disclosure of that information? What is the potential for harm to the relationship in which the information was generated (that is, is there any confidential relationship involving broader societal policies protecting those relationships)? What safeguards can be put in place to protect the information in electronic form in the courthouse or remotely? What procedures and criteria should there be for allowing access in specific instances to otherwise confidential or private information, or in specific instances protecting sensitive or private information when it would

97. See REPORT, supra note 75, at 2.
98. See supra note 43 and accompanying text.
otherwise be publicly accessible? And finally, what is the need for access of the public and the press?

III. CONCLUSION

This is obviously only a beginning sketch. The problem requires a great deal more reflection and thought. What is clear, however, is the need for more focus on the concrete ways in which information in judicial records is being used in our society. The general principles to be used in establishing the right balance are well established. There should be a general presumption in favor of public access, especially when it ensures judicial accountability and facilitates the administration of justice. However, courts must take steps to restrict public access when indiscriminate access conflicts with the administration of justice, when it unnecessarily causes invasions of privacy, or when it exposes litigants, witnesses, and other innocent third parties to threats from the potential misuse of their personal information. The courts should also frame their rules of procedure—local and national—and their instructions to attorneys and litigants to ensure that attorneys and unrepresented litigants know how to take responsibility for identifying sensitive personal information that should be candidates for protection, and are capable of doing so.99 The courts can facilitate this process by encouraging the use of protective orders and motions to seal.

Ironically, it appears that a technological revolution that was supposed to be labor saving will require greater exertion than before from courts and attorneys. In the face of the difficulty of this project, there is a strong temptation to adopt the “public is public” approach, dismissing the concept of “practical obscurity” altogether. However, the temptation to permit indiscriminate electronic public access to court records should be resisted, for it threatens to eviscerate years of careful judicial labor in which courts struggled to achieve an appropriate balance between the competing goals of public access and privacy. In this vein, rejection of the values underlying the concept of “practical obscurity” can be seen as nothing more than a policy decision by the courts to shift the risk of harm to litigants, witnesses, victims, jurors, and

99. For an excellent analysis of the potential malpractice liability of attorneys for failure to take reasonable care to protect their clients’ sensitive personal information in the context of online court filings, see Michael Caughey, Comment, Keeping Attorneys from Trashing Identities: Malpractice as Backstop Protection for Clients Under the United States Judicial Conference’s Policy on Electronic Court Records, 79 WASH. L. REV. 407 (2004).
other innocent third parties whose personal information becomes implicated in judicial proceedings.

This Article has discussed how the use of electronic information makes it much more difficult to balance the need to maintain judicial accountability against the need to protect personal information in court files. As we struggle to maintain this balance, we may discover that the problem is not limited to what is filed in the clerk’s office and that we ultimately have to see this problem in the context of the more general issue of how to regulate the flow of information throughout the judicial system. Traditionally, information exchanged between attorney and client received presumptive protection, while information contained in court files was presumptively open. As electronic information flows from clients to their lawyers and into online court filings, we may be forced to modify this traditional presumption in order to recognize the fact that different types of information have different potentials for misuse depending on their different legal contexts. Even now, the decision of the Committee to keep electronic criminal records off-line marks an important change in our understanding of the need to handle “public” information in court files responsibly. Given the obvious dangers of placing information in criminal case files online, the federal courts were simply forced to treat this information differently from how they treat information in civil files. As more and more courts go online, we may be forced to adopt different rules for access to information in different types of legal proceedings, depending on the potential for that information to be misused.

We have lived in a very forgiving world. The “practical obscurity” of paper judicial records largely sheltered us from the danger of information misuse, while we prided ourselves on our “public” judicial system. The world of electronic information is a far less forgiving place. It is now forcing us to recognize—by our actions, if not yet by our words—that the simple abstract rules developed for a world of paper-based information may no longer suffice to resolve the complex problems of judicial information management. Courts have traditionally been vigilant in protecting individuals from the misuse of sensitive personal information. They must now rise to the difficult task of designing rules to protect litigants and third parties from cyber-mischief and victimization. The failure of the legal system to maintain the ancient balance between access and privacy will lead to the greatest danger of all—inhibiting citizens from participating in the public judicial system. The world of cyber-justice should not be permitted to degenerate into a
Online Court Records

world where victims of crimes are reluctant to come forward; where people are more unwilling to be witnesses or jurors; and where the rich can seek out private judicial forums to resolve their disputes, while the poor and middle classes are faced with an impossible choice—either foregoing justice to maintain their privacy and security; or permitting their sensitive personal information to be commercialized or stolen, and allowing the intimate details of their personal lives to be made available all over the Internet.
THE END OF TECHNOLOGY: A POLEMIC

Louis E. Wolcher*

Abstract: This essay is a philosophical polemic against the essence of modern technology. The piece does not advance a Luddite’s agenda, however, since it describes modern technology’s essence as technological thinking, rather than as the manifold of technical instruments and processes. Technological thinking is not just careful planning towards well thought-out ends. Rather, it is an entire orientation to life, and as such it is a monstrosity: it relentlessly and heartlessly transforms the world’s beings, including human beings, into measurable units of production and consumption that are constantly being judged for their contributions to “productivity.” Nature is thus made into a vast warehouse, and humanity into a standing reserve of “human resources.” Absent from technological thinking is any reflection on technology’s end, in the sense of its ultimate purpose. A synthesis of the thinking of Heidegger, Marcuse, Weber, and Sorel, this essay claims that the ultimate end of technology as such is, and ought to be, freedom for responsibility, and that freedom from necessity is both a condition and a consequence of this. It argues that there is a desperate need for thought and speech to break with technological thinking, and to begin bringing the means of modern technology into contact with its ultimate end.

“The sadness of nature makes her mute.”—Walter Benjamin¹

I. THE END OF TECHNOLOGY: ASKING THE QUESTION

What is the end of technology? Although this question does not mean to ask when technology will end, in the sense of ceasing to exist, it nonetheless remains ambiguous in an interesting and productive way. Consider the “end of technology” understood as technology’s limit. On this reading, the question asks for a determination of the logical endpoint of technology: the limit, or boundary, that circumscribes technology as a concept and allows it to stand forth in thought and speech as a comprehensible whole. Since it is impossible to comprehend a phenomenon like technology in this way without paying attention to its social context, this means that technology’s logical end is inextricably linked to its end in another sense: purpose. When we notice a thing we tend to notice it as something. We notice a telephone as equipment for

* Professor of Law, University of Washington School of Law, William H. Gates Hall, Box 353020, Seattle, Washington 98195. E-mail address: wolcher@u.washington.edu.

calling someone, a house as a place for humans to dwell, a pencil as a utensil to write with, and so forth. What a thing is, and what a thing is for, comprise two sides of the same coin. Thus, it would appear that the purpose of technology, together with a determination of what technology is, gives critical reason an important and singularly synthetic question to think: What is the end of technology?

We live in an age that has been aptly called the “second industrial revolution.” Unlike the first one, the second industrial revolution is characterized primarily by the fact that scientific discoveries are translated routinely and almost immediately into new procedures of production and distribution—new ways to dominate nature’s beings and to control human behavior. Domination and control are not necessarily bad things to strive for, of course. As Herbert Marcuse observes, the instruments of technology by themselves “can promote authoritarianism as well as liberty, scarcity as well as abundance, the extension as well as the abolition of toil.” As concepts, however, domination and control do shed some light on the meaning of our question. Sadism aside, the phenomena of domination and control belong to the category of means rather than the category of ultimate ends. To inquire about the end of technology is thus to ask for it to be determined in such a manner that we can understand what Aristotle called its final cause: its purpose for being what it is. In short, what, ultimately, is technology aiming at?

Technology’s ultimate end is not the same as the sum of its technical performances. Marcuse rightly distinguishes the paraphernalia of modern technology (“technics”) from the totality of technics and their modes of organizing and changing social relationships, prevailing thoughts, and behavior patterns (“technology”). If one were to say that the purpose of the telephone is to speak with people at a distance, that of the Internet to provide instantaneous access to information, and that of the automobile to travel rapidly from here to there, all of these particular ends would remain merely intermediate. One is still entitled to ask what the ultimate point is of having instruments that allow us to speak with people at a distance, gain instantaneous access to information, and make

5. See MARCUSE, supra note 3, at 41.
The End of Technology

all of our journeys rapidly. In technology, as in other things, the most obvious truth is often the least visible: the question concerning the meaning and justification of technology is not itself a technological question.6

We are used to calculating in terms of intermediate ends, and only rarely do we ever really think about ultimate ends. Indeed, this allocation (or misallocation) of our mental energies may be one of the prime effects of modern technology, which tends towards the production of what Horkheimer and Adorno call “purposiveness without purpose.”7 Getting an education that teaches useful job-skills, earning a living, becoming successful by conventional measures of income and status, and so forth—we must not let the self-evident desirability of these ends obscure the fact that they are all merely intermediate. The chain of intermediate ends in a technological society is extremely long, as anyone who has ever spent much time toiling in a modern work cubicle knows. But in the final analysis, they are (or should be) all classified on the level of means. And while Heidegger undoubtedly is correct that the most secure and comfortable path is to make something harmless and insignificant by calling it self-evident,8 respect for the capacity to think requires us to conclude that the question where these intermediate ends ultimately lead, as means, remains wide open. However rational we may be in pursuing a “useful” education, a “good” job, and conventional “success,” these means do not necessarily bring us happiness or lead to the plausible ultimate end of living a good life.

Max Weber’s notorious “iron cage”9 vividly represents the radical difference between intermediate and ultimate ends in the conditions of a technological society like ours. According to the most profound sense of this image, the universal rational pursuit of intermediate ends in such a society (sometimes called the “rat race”) can construct a kind of prison

9. MAX WEBER, THE PROTESTANT ETHIC AND THE SPIRIT OF CAPITALISM 181 (Talcott Parsons trans., 1958). We actually owe the phrase “iron cage” to Parsons, who used these words to translate Weber’s German expression, ein stahlhartes Gehäuse. The latter can also be translated as “a casing, or housing, as hard as steel.” See DEREK SAYER, CAPITALISM AND MODERNITY: AN EXCURSUS ON MARX AND WEBER 144 (1991).
in which the formulation and attainment of any rational ultimate end for the individual becomes difficult or impossible. Perhaps this explains why the couches of psychoanalysts in the West are full of people who have worked very hard to have it all, outwardly speaking, but who inwardly experience their lives as trivial and devoid of meaning. Even retirement, which used to represent the dream of rest, recreation, and the pursuit of personal fulfillment after a lifetime of toil, is under attack from the standpoint of intermediate ends. Consider that a recent article in a prominent magazine asks, rhetorically, “Why should people living longer, healthier lives also feel they are automatically entitled to spend less time working and ever more time in retirement?”—as if the ultimate point of life is to live in order to work, rather than just the opposite.

Weber’s concept of the iron cage also reproaches the relation between intermediate and ultimate ends when it is considered from the global standpoint of humanity as a whole. According to a recent article on the regrettable persistence of hunger in the world, the World Bank’s chief economist, Nicholas Stern, has pointed out that the average European cow receives $2.50 in subsidies, while seventy-five percent of the people in Africa get by on less than $2 in global development aid. No convincing account of some deep and irreconcilable conflict between ultimate ends explains why Europe gives more money to cows than to starving Africans, especially since it is well known that agricultural subsidies in rich countries actually undercut the efficacy of development aid to poor countries by depressing the price of food on world markets. Of course, the pursuit of short-term self-interest by European farmers and their lobbyists is understandable from the standpoint of their intermediate end of maintaining a high standard of living. But a larger perspective gives thought the warrant to ask how this state of affairs could ever be seen as a plausible ultimate end for modern technology, especially given the enormous unrealized productive power that today’s humans have at their command.

With all due respect to Martin Heidegger, who tended to disparage the

12. Id. (“These subsidies also depress commodity prices, undercutting the ability of developing nations to compete in world markets and get their nations off the dole.”).
The End of Technology

language of means and ends, the question of technology’s ultimate end is the right one to ask because it self-consciously appropriates the terminology of technological thinking (the means-end relationship) in order to question technology on its own terms. If there is something about the essence of modern technology that has inflicted a wound on humanity, then it alone is the sword that can heal that wound. For the essence of modern technology is not something abstract and distant from us—on the contrary, it is in us and around us in the form of our world. In an operational and result-oriented world like ours, a fact is the projection of a method for finding it, and a method is the projection of a human purpose. This double insight into the intimate relation between facts and purposes, Is and Ought, gives critical reason the chance to disentangle the terms of the relation, interrogate the mostly hidden purposes of modern technology for their meaning, and yes, also assess modern technology’s ultimate rationality.

The questionability of technology’s contribution to ultimate ends is an ancient theme. In the *Charmides*, one of his earliest dialogues, Plato describes Socrates as expressing considerable doubt about the ultimate end of technological progress. After discussing with Critias the meaning of the kind of wisdom that knows where its own knowledge ends, and hence where its ignorance begins, Socrates relates a troubling dream that he had. The dream concerns a society where everything is extremely well ordered, technologically speaking. Unable to tell whether the dream came through “the horn or the ivory gate” (that is, whether its content was bad and false, or good and true) Socrates goes on to describe what he dreamt:

The dream is this. Let us suppose that wisdom is such as we are now defining, and that she has absolute sway over us. Then, each action will be done according to the arts or sciences, and no one professing to be a pilot when he is not, no physician or general or anyone else pretending to know matters of which he is ignorant, will deceive or elude us. Our health will be

13. MARTIN HEIDEGGER, *HÖLDERLIN’S HYMN “THE ISTER”* 45 (William McNeill & Julia Davis trans., 1996) (“The question as to whether modern technology is a means or an end is erroneous already as a question, because it utterly fails to grasp the essence of modern technology.”).
16. *Id.* at 119.
improved; our safety at sea, and also in battle, will be assured; our coats and shoes, and all other instruments and implements will be skillfully made, because the workmen will be good and true. Aye, and if you please, you may suppose that prophecy will be a real knowledge of the future, and will be under the control of wisdom, who will deter deceivers and set up the true prophets in their place as the revealers of the future. Now I quite agree that mankind, thus provided, would live and act according to knowledge, for wisdom would watch and prevent ignorance from intruding on us in our work. But whether by acting according to knowledge we shall act well and be happy, my dear Critias—this is a point which we have not yet been able to determine.\textsuperscript{17}

Here, at the beginning of Western philosophy, we witness a mind troubled by the relation between technical efficiency and ultimate human ends: a mind willing to question that relation.

Socrates knew that so long as we question something exclusively in terms of what it immediately provides us in comfort or material well-being, we are thinking technologically, not philosophically. Accordingly, the italicized words in the previous paragraph indicate that Socrates was unsure how (or even whether) science and technology, viewed as means, contribute to the ultimate end of living a good life. In this kind of questioning he is not alone in the history of Western philosophy. Consider the brilliant twentieth century philosopher Ludwig Wittgenstein, who expressed his own doubts on the question rather more succinctly: “We feel that even when all possible scientific questions have been answered, the problems of life remain completely untouched.”\textsuperscript{18}

If a Luddite is defined as someone who yearns to shatter the instruments of technology so as to return society to a supposedly pre-technological Eden,\textsuperscript{19} then it must be said that neither Socrates nor Wittgenstein were Luddites. They were thinkers trying to get beneath the glitter of technological progress to reach what is primordial about it.

\textsuperscript{17} Id.


\textsuperscript{19} The Luddites were a group of British workers who rioted and destroyed textile machinery between 1811 and 1816, on the belief that it would diminish employment. The term probably comes from the name of a worker, Ned Ludd, who destroyed stocking frames in eighteenth century England. THE AMERICAN HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE 774 (William Morris ed., 1969).
The End of Technology

Socrates’ case, the primordial is technology’s apparent indifference (or irrelevance) to what human beings ought to do with their lives, as opposed to how they should do what they happen to have decided to do. In Wittgenstein’s case, the primordial is a function of science’s commitment to methodological rigor: the scientific need to predetermine the realm of questions that can be sensibly put to the world also narrows the range of permissible answers, thus ensuring that none of these questions and answers will ever touch what he calls the ultimate “sense” and “value” of the world.20

A question like “What is the end of technology?” thus needs to be thought down to its roots before any decently thoughtful answer to it can be attempted. Most of the time we tend to leap over what is simple and original, and get hung up on the complicated and derivative. And so it is with technology: we tend to leap immediately into seemingly intractable political controversies like technological progress versus the threats that it poses to our privacy, or the preconditions for inducing investment in future technology versus the needs of the poor, in undeveloped and developed regions alike, to enjoy the benefits of present technology. These problems are admittedly pressing and difficult, but they stand no chance of being solved, or even properly understood, so long as the question that grounds them remains unasked. Preceding all questions about particular aspects of technology (including the manifold that is sometimes called “law and technology”) is this one: What is technology?

As a grounding question, the question just asked does not merely seek to uncover correct information about technology’s instrumentalities and support institutions. Before thinking about personal computers, cell phones, global positioning satellites, and “smart” bombs—before inquiring into the laws of patents, copyrights, trademarks, and unfair competition, which institutionally serve modern technology’s tendency to colonize ever-greater spheres of social life by transforming its instruments into saleable commodities—before any of this, thought should endeavor to grasp technology’s essence.

20. WITTGENSTEIN, supra note 18, ¶ 6.41, at 71. When all possible scientific questions have been answered, Wittgenstein contends, “there are then no questions left, and this itself is the answer.” Id. ¶ 6.52, at 73. What is it that the absence of all questions answers? According to Wittgenstein, it answers why language itself cannot solve the most important problems of life: “Is not this the reason why those who have found after a long period of doubt that the sense of life became clear to them have then been unable to say what constituted that sense?” Id. ¶ 6.521, at 73.
II. THE ESSENCE OF MODERN TECHNOLOGY

The essence of modern technology is not what medieval philosophers would have called technology’s *quidditas*, or whatness. Hence, our problem is not to define the word technology, for such a definition would only give us the illusion of a secure and stable content, corresponding to what Heidegger would call technology’s “ontical”\(^{21}\) determination. Before an entity is defined as *what* it is (the ontical), there always prevails an ontological presupposition, on the part of the definer, of what constitutes the being of the entity. In other words, you have to already know what kinds of things count as really existing, as well as the manner of their existing, before you can go around noticing and describing this or that particular existing thing. For example, a Platonist mathematician will insist that the number \(\pi\) really exists on the number line, whereas a philosopher of mathematics may insist that \(\pi\) does not exist, except in the form of a rule that is useful in making certain human calculations.\(^{22}\) Heidegger succinctly (if rather abstractly) describes this insight as follows:

> Whatever is discloses itself only on the grounds of a preliminary (although not explicitly known), preconceptual understanding of what and how such a being is. Every ontic interpretation operates on the basis, at first and for the most part concealed, of an ontology.\(^{23}\)

The difference between what a thing is and how it is (or between a being and the being of that being) invites reflection about the pre-understanding that thought always brings to the project of defining. In a phrase, the ontological difference allows us to see that whatness presupposes howness.

The manner of technology’s existence is thus prior to, and more primordial than, its essence in the ontical sense. The simple “whatness” of a thing is the product of a conceptual system that is reified, as Theodor Adorno puts it, “on the model of a functioning apparatus”: such a system congeals its object into something solid that exists in a time

\(^{21}\) M A RTIN HEIDEGGER, BEING AND TIME 31–35 (John Macquarrie & Edward Robinson trans., 1962) (describing the “ontological difference” that we are elucidating in the text).


that has also been solidified into a uniform series of present moments. However useful this process may be to the technological manipulation of the object-world, it is fatal to critical thought about technology, for it ignores the dimension of historical, lived time that must be accounted for if technology’s ultimate end is to come into view. The point is not that “existence precedes essence,” as Jean-Paul Sartre once said. The point is that a being’s manner of existing is its essence, at least for our purposes. Thus, “essence” in the context of the present investigation can only name the manner in which technology comes to presence and abides through time: essence in the active sense of “essencing.”

Considered dynamically and historically as the ongoing creation of a world in which humans must dwell, modern technology does not just reveal new realms or aspects of beings and articulate how they may be manipulated to our advantage. Such a point of view on technology naively imagines that there exist pre-made humans, on the one hand, and then, on the other hand, instruments of technology that human beings simply use or forbear from using. Technical methods are never just procedures for treating objects: they represent a decision, and a presupposition, about what constitutes the very objectivity of the object world. More important than any use or project, or than any particular technical performance, the essence of modern technology is such that it has radically transformed our relationship to the real. The essence of modern technology is in a very important sense pre-technological, for it co-determines the attitudes, thought processes, and behaviors of human beings before they ever set out to use any of technology’s instruments. It is a form of “challenging,” as Heidegger says: the essence of modern technology “enframes” the world and challenges human beings to assemble and order it in a particular way.

In his most important text on technology, Die Frage nach der Technik (“The Question Concerning Technology”), Heidegger argues that the essence of modern technology (in the sense just described) poses the

27. Heidegger, supra note 6, at 19.
28. Id. at 3–35.
greatest of dangers to humanity and the world.29 Most Cassandras of modern technology utter warnings about the catastrophic potentialities of its instruments: atomic weaponry, bioengineering, the clearcutting of rain forests, and so forth.30 Contradicting this usual account of technology’s dangers, Heidegger says that the threat to humans “does not come in the first instance from the potentially lethal machines and apparatus of technology.”31 He declines to stir up a “demonry of technology”32 by decrying the destructive power of its instruments. Rather, he attempts to warn us that the real danger to human beings comes from an essence of technology that takes the form of technological thinking. The problem is not that human beings employ the means of modern technology, but that they have surrendered their critical reason to the mentality that has brought those means into being: the mentality of Homo faber.

Technology in all ages is never merely a means, Heidegger says: it is a way of revealing the real.33 Men and women do not just use modern technology—they are also used by it. Working all day in front of a computer screen, in a factory, or in a store selling things that people may or may not need, the subject of technology comes home to sit, transfixed, before a television set that constantly bombards its viewers with carefully constructed advertising images, sound-bytes of “news” and “commentary,” and reality shows full of attractive young men and women willing to humiliate themselves for cash. Like a sponge, the subject of technology gradually soaks up what the system makes available in the object-world by way of occupations, news, and entertainment, and in the process undergoes a metamorphosis. Under these conditions the subject “become[s] an object of its object”34 as the Chilean philosopher Rolando Gaete puts it, in a dialectical movement that goes from subject to object-world and object-world back to subject. Although modern social theory specifies that the subject-object dialectic

29. Id. at 26–28.
30. See, for example, Einstein’s warning, in 1950, that the nuclear arms race makes the “fate of humanity” hang in the balance “more truly now than at any known time.” ALBERT EINSTEIN, IDEAS AND OPINIONS 161 (1954).
31. HEIDEGGER, supra note 6, at 28.
32. Id.
33. Id. at 12.
The End of Technology

characterizes any and all human societies, there is something uniquely de-ratiocinating about this one. For this dialectic transforms the ancient values of logic and critical reason into their modern doppelgangers: logistics and calculation. A sort of technological philistinism takes hold, according to which all knowledge and performances must be evaluated exclusively by their immediate utility and their contribution to material values.

The essence of modern technology sets upon nature and challenges it to become a constant source of supply: a “standing reserve” (Bestand), as Heidegger puts it. In this state of affairs the realm of production never stops: what is produced reenters technology immediately as the means of further production. Production becomes production for its own sake, with the result that we consume in order to produce, and live in order to work. This mode of revealing the real is qualitatively different from the ways of thinking that were associated with pre-modern modes of technology. The essence of modern technology manifests what Heidegger calls the “will to planetary ordering,” and it is unique in the dimension of its relentlessness and its tendency to colonize everything and everyone: it never rests. We feel the need to be rational and productive from the cradle to the grave. Even when it is on vacation, technological thinking is busy organizing the environment to extract the maximum amount of “fun” in the available time, recording it all on a digital camera for future moments of enjoyment in the carefully managed free time that it constructs for itself in the interstices of the workday.

From a Heideggerian perspective, today’s world has become a standing reserve because technological thinking allows nature to reveal itself only in the form of what can be continuously computed and counted on for present or future use by human beings. “[E]verything is ordered to stand by, to be immediately at hand, indeed to stand there just so that it may be on call for further ordering.” The idea of “enlightenment” in such conditions becomes the monstrous demand that everything be explained in terms of grounds that are readily accessible to

36. Heidegger, supra note 6, at 17.
37. Heidegger, supra note 13, at 48.
38. Heidegger, supra note 6, at 17.
procedures of domination and control. Thus, technological thinking is not just a reflection of the will to master things—it is also the compulsion to master things on the basis of correct procedure: it leads to what Slavoj Žižek calls “the all-pervasive predominance of ‘instrumental Reason,’ of the bureaucratization and instrumentalization of our life-world.” The chaotic, surprising, and spontaneous in nature and in human relations show themselves as suspicious and dangerous to technological thinking: “problems” to be overcome by appropriate administrative solutions.

One important solution is cooption: the process by which opposition (especially artistic) to the prevailing form of life is absorbed and rendered harmless by the system it indicts. Art at its best has always performed a transcendent and oppositional function vis-à-vis the given reality. However, the members of today’s would-be avant garde are barely able to get ahead of the curve before the curve, constantly on the lookout for something new and exciting to incorporate and commercialize, catches up to them and turns them into “celebrities” and their projects into moneymakers. Marcuse refers to the example of so-called “revolutionary art” like Picasso’s Guernica becoming fashionable and classical—“a cherished museum piece.” One could also mention the disgusting commercialization, in movies and on television, of Anne Frank and the Holocaust, as well as the transformation of the clothing and symbols of 1960s counter-culture into fashionable “retro” statements that can be worn for fun to the office or to the next hippy theme party.

In a world where productivity and efficiency become the measures of all things, artistic images and new ideas can (and do) pass almost effortlessly into slick advertising, political spots, and other forms of self-promotion: think of Gershwin’s Rhapsody in Blue getting us in the right mood to buy airline tickets, images drawn from George Orwell’s 1984 being appropriated to launch a computer company, the aging former film rebels Orson Welles and Dennis Hopper hawking wine and computer

39. Attributing the origin of this point of view on nature to Cartesianism, Richard Watson characterizes its root premise as follows: “Since there is nothing spiritual or sacred in nature, the entire value of the natural world resides in its utility for humankind.” RICHARD WATSON, COGITO, ERGO SUM: THE LIFE OF RENÉ DESCARTES 17 (2002).

40. SLAVOJ ŽIŽEK, THE TICKLISH SUBJECT: THE ABSENT CENTRE OF POLITICAL ONTOLOGY 221 (1999) (mentioning this phenomenon in the context of resistance to it from both the left and the right of the political spectrum).

41. MARCUSE, supra note 3, at 201.

42. GEORGE ORWELL, NINETEEN EIGHTY-FOUR (1949).
The End of Technology

solutions for business on television, and ExxonMobil’s “proud” sponsorship of Masterpiece Theater on Public Television. Under such conditions, artistic images of transcendence are invalidated by their incorporation into the daily reality of production and consumption, and the artist’s “Great Refusal” — the protest against that which is — is transformed into a Great Affirmation of the latest (un)subversive fashions, modes of transportation, computer games, alcoholic beverages, and television programs. And lest art’s insidious and latent affirmation of the commercial world gets lost on the artists themselves, copyright law is there to help them transform their simulacra of protest into profitable exchange values.

Thanks to what Marcuse calls the “frightful science of human relations,” technological thinking knows just what buttons to push in people, which instincts (primarily greed, sex, violence, and sentimentality) to pander to and gratify in order to lead the consumer docilely into the chute. This cynical practice of “repressive desublimation” lets loose instincts that used to be strongly repressed by traditional law and social conventions, only to reroute their profitable satisfaction by business into a dialectic of stimulus and response that binds the consumer ever closer to the system. As Marcuse notes, the process of creating repressive needs bears a strong resemblance to political totalitarianism: just as the latter attempts to coordinate all social performances through political control, the technological kind of totalitarianism coordinates social behavior through the manipulation of needs and aspirations by vested interests. Think of $10,000 watches that keep the same time as a $20 Timex; $1,000 designer handbags and briefcases that hold the same amount of material as durable bags costing a tenth as much; humungous SUVs that transport you with half the fuel efficiency, and at twice the price, of smaller cars; and $2,000 “high end” shoes that are harder on your feet than a good pair of cheap sneakers. The scary mass consciousness of political totalitarianism finds its analog in the ubiquitous and only slightly less scary “happy consciousness” of technological totalitarianism. This is the carefully nurtured “belief that

44. Id. at 63.
45. MARCUSE, supra note 2, at 118.
46. Douglas Kellner, Introduction to MARCUSE, supra note 3, at 1, 11.
47. MARCUSE, supra note 43, at 3.
the real is rational and that the system delivers the goods,48 which makes it well-nigh impossible for the technological subject to break the spell of the things that are by naming the things that are absent in his life:49 like a pacified struggle for existence, or freedom from the planned obsolescence that keeps him working overtime to replace things that are already outmoded almost the day they leave the store, or time to be free instead of acting like the wage-slave of his expensive wristwatch.

Technological totalitarianism is linked to a radical transformation of the meaning of leisure time. The good news is that more people than ever have it; the bad news is what it has become. Traditionally conceived as time full of freedom for the world, nature and culture for those who were lucky enough to enjoy it, leisure time is now a continuum of blank, vacant time that the workday produces as a sort of by-product. The technological subject feels the need to fill this vacant time by the acquisition and passive consumption of commodified units of “entertainment” that someone is always happy to sell, thereby enabling people to while away the time until the beginning of the next workday. Even the great national parks become structured environments, with well-marked paths and learning centers that are suitable for the work-like task of “enjoying nature”—all under the supervision of a cadre of park rangers whose job is to police them for the kinds of uses that the system deems inappropriate. In the terminology of Michel Foucault, technological thinking is making ever-larger swaths of nature into “panopticons,”50 where increasingly refined procedures for monitoring and knowing are creating ever-increasing opportunities for discipline and artificial needs-creation in the service of the system. The end result is a depressingly “comfortable, smooth, reasonable, democratic unfreedom”51 that compensates individuals for the alienation they experience during their work lives by giving them prefabricated and carefully engineered opportunities for titillation in their vacant time. The cost of this titillation is not just the capital and social labor necessary to produce it, but also the hidden cost of needs manipulation by the system: the technological subject must now work more hours, and incur more debt, to be able to afford the most gargantuan SUV, the latest version of the Xbox, and annual trips to the theme park, where history and reality

48. Id. at 84.
49. Id. at 68.
are reduced to predigested and wholesome bites of “fun.”

Marcuse’s sober conclusion about the meaning of all this warrants more than a little reflection: “Free choice among a wide variety of goods and services does not signify freedom if these goods and services sustain social controls over a life of toil and fear.” 52 Although the distinction between what is needed and what is merely wanted has always been historically determined, technological thinking does not make it easy for the subjects of technology to draw their own informed conclusions about the difference between true and false needs. If, with Marcuse, we define false needs as those that “perpetuate toil, aggressiveness, misery, and injustice,” 53 then critical thought requires a certain distance, or dissociation, from what it strives to comprehend if it is to distinguish the false from the true in the realm of needs. But distance from the given reality is precisely what technological thinking tends to deny to its subjects, with the result that individuals in prosperous countries instinctively feel the need to work harder and harder to afford more and more frou-frou. This is not greed, or at least it is not just greed. This is sheer madness, especially when one stops to consider that the finite time available to each individual is their irreplaceable and ever-dwindling font of life itself, and that we live in a world in which hundreds of millions are not even getting their most vital needs met. 54

The essence of technology has not always shown itself in these ways. The word technology comes from technē, which to the Greeks meant the human art of revealing what is concealed. 55 Technē is the knowledge that belongs to the activity of poiēsis, which means the production and bringing-forth of something. 56 As Heidegger notes, the word technē never named the instruments and material apparatus of technology: it always represented the human know-how to transform the beings of nature (physis) into new forms. 57 Nature tends to conceal itself, said Heraclitus, 58 and technē originally signified all modes of revealing what nature hides. The homemaker, the artist, the craftsman, the politician,

52. Id. at 7–8 (emphasis added).
53. Id. at 5.
54. With Marcuse, we will define vital “needs” as “nourishment, clothing, lodging at the attainable level of culture.” Id.
55. HEIDEGGER, supra note 6, at 12–13.
56. Id. at 13.
57. See id. at 12–13.
58. HERACLITUS, FRAGMENTS 71 (T.M. Robinson trans., 1987) (Fragment 123).
and the lawyer: each possesses its own particular kind of technē.\textsuperscript{59} Within all of its many spheres of operation, the Greeks thought that human technē transformed nature as a whole into particular beings that display their own proper order in the world of human needs and aspirations. However, the ordering effected by the Greek technē eventually came to rest, whereupon physis (nature) continued to reveal and produce itself from out of itself (more on this later). Production (poiēsis), for the Greeks, was first and foremost nature’s production of itself, on the basis of which human beings could then produce things by means of a different kind of production: technē.\textsuperscript{60}

The case is otherwise with the standing reserve that modern technology has made out of nature: nature’s self-production is hidden, while the earth shows itself in the form of a huge warehouse full of “resources” that just lie there, waiting to be counted and then exploited, or else saved for future use. While technē in the Greek sense possesses the object of production only until it is finished, thereafter releasing it into the realm of the non-technological,\textsuperscript{61} modern technology never once loosens its grip on the object-world. Here are some examples, which we will place behind bullet-points so that the (odious) technological form will mimic the content:

- The soil becomes a standing-reserve of crops, minerals, and energy to be administered by agribusiness and the energy industry. One no longer runs one’s fingers aimlessly or reverently through the loamy soil—one monitors the dirt for its capacity to grow crops and produce ore.

- The rivers become sites for the construction of dams, and flowing cisterns that are used for irrigating vast farmlands and filling reservoirs in sprawling cities and suburbs. What is left over is carefully husbanded so that salmon can find a place to breed for the fishing and restaurant industries, and so that city people can go on guided (and usually expensive) river-rafting tours.

- The air becomes a standing-reserve of oxygen for industry, and then breathable gas to be monitored for “air quality.” No longer the

\textsuperscript{59} \textsc{Heidegger, supra} note 6, at 13 (noting that technē is not just the art of the craftsman, but also “the arts of the mind and the fine arts”).

\textsuperscript{60} \textit{Id.} at 10–13.

\textsuperscript{61} \textsc{Martin Heidegger}, \textsc{Plato’s Sophist} 29–30 (Richard Rojcewicz & André Schuwer trans., 1997).
dome of heaven, the sky also becomes an indifferent medium through which airplanes and rockets can fly.

- The sea becomes a standing reserve of “marine resources” to be harvested by the fishing and whaling industries, gawked at by snorkeling tourists, and filmed for couch potatoes to watch in nature programs on television. No longer Homer’s mysterious and awesome “blue girdler of the islands,” the ocean becomes a kind of stage set through which vacation cruises sail, and in which suntanned tourists sport and play when they are on vacation at the shore.

- The shrinking patches of primordial nature left to us become a standing-reserve of potential “experiences” to be packaged and sold as McNature by the leisure and tourist industries. Nature as a consumer good thus tends to shove nature as a source of unbought wonder off the shelf. Even conservation is sold as being “good for business,” further confirming nature’s diminishment to the status of an exchange value.

- Life becomes a realm of “bio-diversity” to be divided up, counted, and assessed for its possible contributions to human welfare. The grand spectacle that is life no longer lives for itself—it lives in order to be useful and productive. If they are lucky, “endangered species” get allotted bio-cubicles in the form of a few hundred protected acres; otherwise, they die out, or are whisked away to zoos so their genotypes can be conserved for future propagation and reinsertion into whatever little bits of their habitat remain.

- Language is also radically transformed. First changing from words that uncover beings into words that represent them, language then morphs from a technique of representation into a productive instrument: a crisp, hip, quasi-mathematical language that efficiently communicates operational commands and responses. Thought follows speech’s suit, honing closely to the immediately useful, and steering away from anything that is even remotely contemplative, speculative, or otherwise “useless.” Meanwhile, the constant drone of background music, in public and private places alike, seems to

63. See HEIDEGGER, supra note 6, at 11–12 (comparing the Greek idea of truth as the uncovering of beings (αληθεία) with the modern concept of truth as the “correctness of an idea”).
64. Cf. WALTER BENJAMIN, Karl Kraus, in BENJAMIN, supra note 1, at 239, 242.
confirm Adorno’s thesis that modern mass music fills the silence that spreads when speech is reduced to silence because of anxiety, work, and undemanding docility.65

- The complex web of human actions and reactions, instincts and motives, becomes a standing-reserve of predigested factoids and disaster news to be “spun” and exploited by politics, the media, and mass advertising. Remember the panicky run on duct tape shortly after the attacks of September 11, and the alleged (but as yet unfound) “weapons of mass destruction” that played such a large role in stirring up popular support for the second war against Iraq? Or how about the phenomenon of worldwide terrorism being reduced to evildoers who “hate our freedom,” whose real motives and grievances we cannot and should not try to understand, and who deserve only to be exterminated in the glorified bug-hunt that high-tech warfare aspires to become?

- The political realm is diminished to the chance the masses have to select their masters every couple of years from a slate of candidates whose campaigns are funded by vested interests, and who invariably cut their cloth to the wind of opinion polls that collect droplets of pre-manipulated “public opinion” like the atmosphere collects moisture. Indeed, the machinery of politics virtually requires that the form of political thought conform to a set configuration of public opinion that can easily be made available on demand to prove or disprove virtually anything about “what the people think”: as if the point of politics were to register knee-jerks instead of engaging widespread public debate on difficult (and often tragic) choices.

- Last, and most ominously, even men and women are transformed into a standing reserve of “human resources,” where “[i]ndividual distinctions in the aptitude, insight and knowledge [of workers] are transformed into different quanta of skill and training, to be coordinated at any time within the common framework of standardized performances.”66 These “human resources” are managed, consumed, and discarded on the basis of what Marcuse calls the “Performance Principle,”67 according to which “everyone

66. MARCUSE, supra note 3, at 44.
67. MARCUSE, supra note 2, at 197.
The End of Technology

has to earn his living in alienating but socially necessary performances, and one’s reward, one’s status in society will be determined by this performance.”68 In a nutshell: the system tells you to shape up or ship out. Faced with such a choice, it is little wonder that so few people do anything other than shape up.

There are counter-tendencies, of course, within a world that has been transformed into a standing reserve. Were this not so, it would be impossible to account for the enduring popular appeal of Willy Loman’s argument to his boss, in Death of a Salesman,69 for why he should not be peremptorily fired for lack of productivity after thirty-four years with the firm: “You can’t eat the orange and throw the peel away—a man is not a piece of fruit!”70 (The argument did not succeed, by the way: Willy was fired.) An embattled labor movement, environmentalism and its various Green parties, the grassroots movement to “downsize” career expectations and material possessions, organized opposition to the most pernicious effects of global capitalism on the environment and on the Third World’s poor, the international human rights movement: these represent some of the counter-tendencies. However, from the standpoint of technological thinking, these counter-tendencies will never be more than individual “preferences” to be managed and controlled through the political process. The essential terms of the debate—cost-benefit analysis, interest group politics, the language of “resources,” “human capital,” and “tradeoffs”—are predetermined by the dominant discourse and by the stranglehold that it exerts on the human imagination. Although this discourse does not necessarily produce “bad results,” it does require thought and politics to imagine the world in purely logistical terms. Logistics in its original sense is the science of procuring, maintaining, and transporting military materiel on command. But in today’s world it has come to characterize the way instrumental reason achieves “results” in a political and social environment that is construed as a battlefield of conflicting interests. Logistical methods marshal and disburse arguments, public opinion, and resources in pretty much the same way a supply sergeant assembles and dispatches bullets and C-rations.

Technological thinking is impossible to evade whether we affirm or deny it. “There is no personal escape,” writes Marcuse, “from the

68. Id.
69. ARTHUR MILLER, DEATH OF A SALESMAN (1949).
70. Id. at 82.
apparatus which has mechanized and standardized the world.” He illustrates this point with a simple example that has lost none of its chilling accuracy in the sixty-two years since it was written; indeed, if anything, the example has gained in accuracy:

A man who travels by automobile to a distant place chooses his route from the highway maps. Towns, lakes and mountains appear as obstacles to be bypassed. The countryside is shaped and organized by the highway: what one finds en route is a byproduct or annex of the highway. Numerous signs and posters tell the traveler what to do and think; they even request his attention to the beauties of nature or the hallmarks of history. Others have done the thinking for him, and perhaps for the better. Convenient parking spaces have been constructed where the broadest and most surprising view is open. Giant advertisements tell him when to stop and find the pause that refreshes. And all of this is indeed for his benefit, safety and comfort; he receives what he wants. Business, technics, human needs and nature are welded together into one rational and expedient mechanism. He will fare best who follows its directions, subordinating his spontaneity to the anonymous wisdom which ordered everything for him.

Better choices, more choices, greater efficiency, a smooth and well-ordered existence: how self-evidently desirable this all sounds to those in thrall to technological thinking; yet how pernicious this all can become for those who value spontaneity and freedom above conformity and docility.

Not only is the man’s behavior in Marcuse’s example perfectly “rational,” but the dissolution of all rationality “into semi-spontaneous reactions to prescribed mechanical norms” means that, from the standpoint of technological thinking, only a crank would insist on freedom of action to make journeys in some other, less efficient, way. The cold, wet blanket of conformity to the “correct” thing to do makes any dissent that dares question its correctness from the standpoint of its ultimate truth look like a symptom of mental illness, rather than the expression of an alternative vision of what being rational means.

Nowhere is this kind of ascendency of means over ends better

71. MARCUSE, supra note 3, at 46.
72. Id.
73. Id.
The End of Technology

illustrated than in the standard neo-conservative argument that free markets are almost always superior to government planning and regulation for the production and distribution of goods and services. While this argument can be shown to be technically correct in a large number of circumstances (although perhaps fewer than is claimed by free market fundamentalists) the point is not which form of social organization is “best.” Since both markets and government planning fall into the category of means, the word “best” in this context signifies a criterion that focuses on the efficiency of the means of production and distribution, rather than on their ultimate end. “Best” merely signifies what Weber calls the most “objectively rational” means to reach some “subjectively rational” end: an end that must be known in advance if the means are to be rightly called either rational or objective.74 But neo-conservatism’s brand of “economic liberalism” interprets the formal liberty to enter into market relationships as an ultimate value regardless of the actual ends being pursued.

The argument that the market rightly leaves the formulation of ultimate ends to the individual rather than to the collectivity ignores the fact that the institution of the market, as well as the multitude of contingent laws that make it what it is at any given place and time, are themselves collective (political) decisions. The laws of property, torts, contracts, unfair competition, and so forth construct the market—they do not merely respond to it from the outside. That little word “entitlement” hides a secret that most free market ideologues would like to keep under wraps: without entitlements, there can be no market in the modern sense of the word, and without government defining and protecting them, there can be no entitlements. Since entitlements are a form of stealth government, this means that even so-called free market capitalism is really socialism by another name. By intensively specifying the basic rules of the game, and standing ready to back them up by force (law), the collectivity thus becomes at least co-responsible for the range of choices that emerge from it. If you would rather write poetry and live within an unspoiled natural world, but feel compelled by economic and social circumstances to get a job and an apartment in a grungy city, this narrowing of your effective choices is no less a collective decision than if it had been centrally planned. But more importantly, if we are right

that technological thinking and the world it creates make the formulation and assessment of ultimate ends extremely difficult for the individual, then the economic liberal’s argument comes down to saying that we should always wait for the ultimate ends of individuals to emerge from conditions that are aggressively hostile to them. This is roughly analogous to Henry Ford’s reputed remark about consumer choice in the days of the Model-T: “you can have any color car you want, so long as it’s black.”

The economic subject of early nineteenth century laissez-faire capitalism lived in a world where one could dream of breaking through the flotsam and jetsam of archaic laws, social conventions and popular superstitions that impeded the journey to productivity and profits. Moreover, the individualistic rationality of the industrial revolution “was born as a critical and oppositional attitude that derived freedom of action from the unrestricted liberty of thought and conscience.”

Today, however, the race for productivity and profits is taken for granted to such a degree that the modern economic subject has difficulty imagining any other way of life, no matter how miserable he is in this one. Going far beyond Karl Marx’s notion of alienation, technological thinking makes private power relations appear “not only as relationships between objective things but also as the rule of rationality itself.” The prospect of standing up to The Man, while never a safe thing to do in any era, shows itself in our age to be the very height of irrationality. After all, if The Man is “delivering the goods,” who but a lunatic would want to question his power? What is more, since just about everyone who works is part of some bureaucracy, whether public or private, this means that just about everyone has some degree of power over others in the bureaucratic chain. In such an environment, one can imagine oneself as both administered and administrator, with all the petty pleasures of domination that the latter role makes available to offset the sting of being dominated in the former role. The result of this “you dominate me, and I’ll dominate her” mentality is that the system as a whole tends towards what Marcuse calls “technical self-administration.” This nearly endless chain of domination-submission-domination would be

75. MARCUSE, supra note 3, at 60.
76. See KARL MARX, The Economic and Philosophical Manuscripts of 1844, in KARL MARX: A READER 35, 37 (Jon Elster ed., 1986) (alienation defined in terms of man’s “loss of the object and bondage to it”).
77. MARCUSE, supra note 3, at 57–58.
78. Id. at 59.
laughable were it not so damned pernicious.

The historical transformation of the individualistic rationality of the nineteenth century into the technological rationality of the late twentieth and early twenty-first centuries represents a profound shift in the meaning of self-interest.\textsuperscript{79} The radical autonomy of the nineteenth century’s pioneer, frontiersman, and swashbuckling entrepreneur has been replaced by people’s dependence on institutional expectations to be fulfilled. No longer is proud autonomy an aspect of self-interest: on the contrary, it becomes an \textit{obstacle} to, and heteronomy becomes a condition of, rational action.\textsuperscript{80} The one who goes his own way, who marches to the beat of a different drummer, who dissents and is not a “team player”: such a one has never had an easy time of it. But from the standpoint of technological thinking he is both more and less than just an annoying eccentric: on the one hand, he is a troublemaker and a nuisance to the smooth functioning of the apparatus, but on the other hand, he is so easily ignored and marginalized that he hardly poses a threat at all. The complainer and the naysayer may reject what Adorno calls the “traditional thinking which pins things down and organizes them in terms of rigid concepts,”\textsuperscript{81} but unlike Adorno they just fade away into obscurity: they do not get the jobs, the company cars, or the bonus checks. As a consequence, today’s maxim of rational self-interest has become, more than ever before: “You have to go along to get along.”

While the pursuit of self-interest has always been conditioned by prevailing social conditions, never before has the individual’s conception of himself and his choices been so \textit{standardized}. If the technological subject is fortunate enough to get a good (and this usually means expensive) education, it is true that he or she can choose from a large menu of careers. But the norms and expectations governing what it means to act, to think, and to consume like, say, a lawyer, a doctor, or a corporate executive are virtually the same everywhere. Today’s law students, for example, are right in fearing that their “true” selves are at risk if they become successful corporate lawyers. Unlike the rough-edged and non-standardized lawyering of Lincoln’s day, the expectations of the “successful lawyer” in today’s world tend to transform people into Lexus-driving, three-piece-suit-wearing, briefcase-

\textsuperscript{79} The terms “individualistic rationality” and “technological rationality” are Marcuse’s. \textit{Id.} at 44.

\textsuperscript{80} \textit{Id.} at 50.

\textsuperscript{81} \textsc{Adorno, supra} note 24, at 15.
carrying caricatures, and legal performances into standardized moves in an elaborate game. Whereas one would like to think that the true degree of human freedom is not just determined by the range of pre-made choices open to the individual (e.g., “casual Fridays” at the law firm), but by the freedom to radically transform what is chosen, as well as by the existence of a residual freedom to opt out of the system altogether. But what one would like to think is not what the system trains us to think: “The countless agencies of mass production and its culture impress standardized behavior on the individual as the only natural, decent, and rational one.”82

In the nineteenth century, individuals could plausibly escape into what Hannah Arendt calls “non-society strata” lying outside the established order.83 In the days when “society” in principle referred to only certain classes of the population, Arendt notes, a person could identify with excluded groups like the proletariat, Jews, and even homosexuals.84 But today, she astutely observes, “[a] good part of the despair of individuals under the conditions of mass society is due to the fact that these avenues of escape are now closed because society has incorporated all strata of the population.”85 In other words, most people in advanced technological societies feel that they are part of the system: they can succeed within the system, or they can fail and become mediocrities. Beyond that, nothing else is even imaginable. The law student who chooses to become a public-interest lawyer, for example, stays within an overall system of lawyering that construes progress as successfully manipulating the doctrinal apparatus. However noble this kind of lawyering is when seen from a position that values lifting up the oppressed, it leaves the essential structure of the legal machine intact. The fact that virtually everyone comes to value legal means (primarily the discourse of individual rights and duties) is all the system needs in order to perpetuate itself.86 For the occasional vindication of the rights of undocumented aliens and welfare recipients fits seamlessly into the perpetual vindication of the rights of government and big business. The

82. HORKHEIMER & ADORNO, supra note 7, at 21.
83. HANNAH ARENDT, BETWEEN PAST AND FUTURE 200 (1968).
84. Id.
85. Id.
86. Cf. Isaac Balbus, Commodity Form and Legal Form: An Essay on the Relative Autonomy of the Law, 11 LAW & SOC’Y REV. 571 (1977) (noting that the valued form of legal rights, based on equal citizenship, is far more important than particular legal results for maintaining the domination and oppression of a class-based society).
The End of Technology

point of mentioning this is not to make the familiar Marxist argument that rights are ideological, anti-radical, and bad. \footnote{See, e.g., Duncan Kennedy, \textit{Antonio Gramsci and the Legal System}, 6 ALSA FORUM 32, 36–37 (1982) (noting that “people like Marx thought of [rights] as the quintessence of bourgeois false consciousness”).} \textit{The point is that even the legal “rebel” seems perversely fated to become part of the overall system of technological thinking.}

Economic liberalism’s vaunted criterion of Pareto optimality\footnote{The principle of Pareto optimality holds that “a configuration is efficient whenever it is impossible to change it so as to make some persons (at least one) better off without at the same time making other persons (at least one) worse off.” \textit{John Rawls, A Theory of Justice} 67 (1971).} is only efficiency in the pursuit of intermediate ends in any society where people are forced to pursue \textit{exclusively} intermediate ends in order to survive. For in such a society, the vast majority of the population express their “preferences” first and foremost in terms of what will get them through to the next paycheck or the next mortgage payment. This behavior is admittedly an objectively rational means to the end of survival, but it is subjectively rational only if people’s ultimate end in life is merely to survive the struggle for existence, like a beast that dies of old age. The criterion of a “Pareto-rational society” under such conditions thus tends to sanctify the status quo, no matter how alienated and threatened people actually feel. The result is the glorification of a kind of non-bloody conflict in the marketplace and in the halls of Congress that corresponds to Foucault’s inversion of Clausewitz’s notorious proposition about war and policy: “politics is the continuation of war by other means.”\footnote{\textit{Michel Foucault, “Society Must Be Defended”} 15 (David Macey trans., 2003). Carl von Clausewitz actually wrote (in translation): “War is a mere continuation of policy by other means.” \textit{Id.} at 21 n.9.} Conservative and mainstream social science thus decide, explicitly or implicitly, that they are \textit{for} conflict, and they refuse to take seriously the possibility that the category “conflict” could be replaced by a peace that is not merely juridical, but also social and economic.\footnote{See \textit{Adorno}, \textit{ supra} note 24, at 68.} One is again left with the impression that the widespread belief that efficient markets always make for the best of all possible worlds represents yet another victory for technological thinking, and a triumph of means over ends.

The effect of technological thinking’s “pessimistic anthropology”\footnote{\textit{Jürgen Habermas, Afterword to Marcuse, supra} note 2, at 238.} on our perception of world-wide social inequality, poverty, disease, hunger, and exclusion is predictable: these travesties show themselves as


88. The principle of Pareto optimality holds that “a configuration is efficient whenever it is impossible to change it so as to make some persons (at least one) better off without at the same time making other persons (at least one) worse off.” \textit{John Rawls, A Theory of Justice} 67 (1971).


90. See \textit{Adorno, supra} note 24, at 68.

91. \textit{Jürgen Habermas, Afterword to Marcuse, supra} note 2, at 238.
regrettable but inevitable facts of nature, rather than as humanly avoidable consequences. The theory that, in a market society, social protectionism and concrete measures aimed at helping the poor usually wind up hurting them in the end is a powerful narcotic, as the thousands of law students who have learned even a little bit of law-and-economics during the past thirty years can confirm. 92 Marx, in his day, called religion the “opium of the people.” 93 To paraphrase the great labor radical Joe Hill, 94 by this Marx meant that religion’s promise of pie in the sky when you die takes people’s minds off the fact that they have to live on hay down below while they are alive. It would seem that the resigned viewpoint that market-based instrumental reason aimed at intermediate ends is the only viable course left to us has become the opium of today’s progressive intelligentsia, especially since the collapse of communism in the Soviet Union and Eastern Europe demonstrated the inherent weaknesses of planned economies. But historical communism’s planned economies were always just as oblivious to ultimate ends as technological thinking in Western democracies. The point is not that markets should be seen as the best (or the least bad) solution to delivering the goods. The point is that the obsession with “delivering the goods” has tragically eclipsed all other considerations in public discourse about the relations among technology, nature, and life.

Julia Kristeva recently asked whether “life as an event and a source of questioning [has] become outmoded because it has been secured, standardized, and trivialized by technology.” 95 But to ask this question is to refute it: the point is, or should be, also to bring technology’s ultimate end into view, so that technological means can be assessed in terms of their qualitative and quantitative progress towards that end. The essence of modern technology, in the form of technological thinking, almost guarantees that the question concerning technology’s ultimate end is not asked; it almost guarantees that technological means remain blind to any end other than their own perpetuation. Despite these tendencies, thought

---

92. For a general analysis of the way law school trains students to accept invidious social hierarchies of all sorts, see DUNCAN KENNEDY, LEGAL EDUCATION AND THE REPRODUCTION OF HIERARCHY (1984).


94. His song The Preacher and the Slave, in INDUSTRIAL WORKERS OF THE WORLD, LITTLE RED SONGBOOK 36 (19th ed. 1923), contains the following lines: “You will eat, bye and bye, / In that glorious land above the sky; / Work and pray, live on hay, / You’ll get pie in the sky when you die.”

95. JULIA KRISTEVA, HANNAH ARENDT 45 (Ross Guberman trans., 2001).
The End of Technology

owes itself, and humanity, the task of discovering whether the intense complexity of the means of modern technology may have covered over a possibly intense simplicity of their ultimate end.

III. MODERN TECHNOLOGY’S ULTIMATE END

Technology’s ultimate end should not be confused with the previously described effects of technological thinking. If, as we have attempted to show, technological thinking lives in forgetfulness of ultimate ends, this does not imply that technology has (or should have) no ultimate end. And if technological thinking has the baleful consequences that we have attributed to it, this does not mean that technology’s ultimate purpose is the dehumanization of men and women and the transformation of nature into a warehouse, any more that the use of a hammer as a murder weapon indicates that murder is (or should be) the ultimate purpose of hammers.

The use of the word “should” in the previous paragraph indicates something important that we have yet to make sufficiently clear: Although it may seem that the effects of a social phenomenon are there for all to see, neither its effects nor its ultimate purpose can come into focus clearly without collapsing the formal distinction between is and ought, fact and value, performance and aspiration. That this has been our guiding thought from the outset is indicated by the proposition that was put forward in the very first paragraph: What a thing is (fact), and what a thing is for (value), comprise two aspects of the same theme. The insight that thought always approaches a thing from a particular value-position is what saves the endeavor to determine the purpose of a whole institutional complex like modern technology from becoming what Lon Fuller calls “naive teleology.” In a sense, all descriptions of what is—"the facts"—are teleological, since they represent what the describer thinks is important and meaningful about the situation. To paraphrase the philosopher Franz Rosenzweig: although a fact is never identical to its classification, it is nonetheless always being judged according to its

96. The first is what philosophy for centuries, following Aristotle, has called technology’s own causa finalis (final cause), while the second is the effect of technology seen as a causa efficiens (efficient cause). See HEIDEGGER, supra note 6, at 6.

97. LON L. FULLER, THE MORALITY OF LAW 146 (1969). The book as a whole articulates and defends Fuller’s thesis that law is a “purposive activity” that cannot be understood merely as a fact. See, e.g., id. at 95–151.

classification.\textsuperscript{99} It follows that any description of, or meditation on, technological thinking necessarily implies some level of satisfaction or dissatisfaction with what it is, as compared to other possibilities for human existence. Since history has given us modern technology, and has delivered us over to its essence, in the form of technological thinking, our task here can only be to comprehend and recognize these facts in terms of what we think they mean for humanity. In articulating technology’s meaning, thought transcends the semiotics of the here-and-now: thought transforms the self-perpetuating semiotics of description into genuine critique. The concept of the ultimate end of technology that we have been developing thus corresponds to Marcuse’s trenchant observation, in One-Dimensional Man, that “[i]n social theory, recognition of facts is critique of facts.”\textsuperscript{100}

From the standpoint of the individual, an ultimate end is the idea of a state or performance the consummation of which has value in itself—not because it leads to some other end, but simply because it is. Kant, for example, famously elevated the human being to the status of an ultimate end in one of his most powerful expressions of the Categorical Imperative: “Act so as to treat man, in your own person as well as in that of anyone else, always as an end, never merely as a means.”\textsuperscript{101} To put it in rhetorical terms, the statement of an ultimate end represents its own justification. Since the whole of an individual’s life is never a project that is completed during his or her lifetime, the ultimate value of a human life as a whole is more difficult for the individual to bring into view. Nevertheless, most individuals would readily agree with Kant that at least their own lives have ultimate value for them. Even in the midst of the rat race that technological thinking has constructed, people do occasionally manage to take time to wonder whether their life as a whole is going anywhere meaningful. The case of individuals attempting to grasp the ultimate value of something that transcends their here-and-

\textsuperscript{99} FRANZ ROSENZWEIG, UNDERSTANDING THE SICK AND THE HEALTHY: A VIEW OF WORLD, MAN, AND GOD 50 (Nahum Glatzer trans., 1999) (“The act is not identical with its classification, but it is judged according to its classification.”).

\textsuperscript{100} MARCUSE, supra note 43, at 118.

\textsuperscript{101} IMMANUEL KANT, Metaphysical Foundations of Morals, in THE PHILOSOPHY OF KANT 154, 195 (Carl J. Friedrich ed., 1993). Walter Benjamin thought that Kant’s categorical imperative had not elevated humanity high enough: “One might, rather, doubt whether this famous demand does not contain too little, that is, whether it is permissible to use, or allow to be used, oneself or another in any respect as a means.” WALTER BENJAMIN, Critique of Violence, in BENJAMIN, supra note 1, at 277, 285 n.*.
The End of Technology

now, and that is also never before them in the form of a whole, gives us a pattern, or precedent, for thinking about the ultimate end of technology. For modern technology, like life itself, also transcends any one individual’s here-and-now. Hence, however difficult it may be to accept the notion that something like technology could itself have an ultimate end, that is precisely the claim this essay makes. We claim that technology does indeed have an ultimate end, something that leads nowhere else, and that constitutes its own justification.

Although death is the termination of life, the ultimate end of life is not death. The “ultimate” in the context of the present investigation is never merely the last, but also always the first. It is the origin of all comportment, of all that is intermediate. Just as the farthest point at the commencement of a round trip journey is always the beginning, so too the ultimate, as origin, is that to which everything secondary eventually returns. Although the individual may desire much, including happiness, there is something that is more original than any happiness, something that puts happiness into question as one of many possible goals. The origin, the ultimate “wherefore,” of all human doing is never this or that condition or state of affairs, for no condition or state of affairs ever abides long enough to be ultimate.

Just as the human being as such cannot rightly be reduced to the merely biological life with which it coincides, so too the ultimate end of technology, as a manifestation of human being-in-the-world, does not coincide with this or that project or performance. If, for heuristic purposes, we follow the Western metaphysical tradition and say that the human being is an embodied soul, this gives us a clue for what we are seeking. We must attempt to identify the “soul” of technology that is embodied in all the countless things that humans do with its instruments. In short, the criterion of such a thing must be its propensity to abide throughout countless performances, in much the same way that the soul is said to abide through the many changes (including deteriorations) that occur in the aging human body. There is only one thing that abides in this way as long as human life lives, and in its constant abiding it is both the origin and the return of all that is properly human: freedom. Let us then say it plainly: The ultimate end of technology is a particular kind of freedom.

A word or two on method: we have both “found” freedom in our meditation on the ultimate end of technology and laid freedom down as an axiom for the assessment and further development of technology. It bears repeating that in social criticism the relation between Is and Ought
is always dialectical. Those who wish to “prove” freedom’s impossibility do not escape it: they merely lay it down in a negated form (unfreedom) as the axiom of their lives. They freely choose to bend their necks to the yoke of technological thinking. We choose otherwise. For us, “freedom” names the possibility of decisively breaking with technological thinking—the possibility of changing history, for better or worse.

If we define technology as such to be technical means in the service of an ultimate end, then we are forced to conclude that the previously described essence of modern technology is at war with freedom as its ultimate end—at war with that of which it is the essence! In this curious state of affairs, it is as if the essence of the human being—animal rationale (“rational animal”), as traditionally determined—were subverting the very purpose and meaning of human life: as if man had finally succeeded in outsmarting himself. For the essence of modern technology, as technological thinking, has no ultimate end. In this dimension its rationality is fundamentally irrational, at least if one takes the viewpoint that the purpose of society as a whole is to preserve and unfetters the people of which it is composed.102

Like all things human, the essence of modern technology makes a world—an odious world, perhaps, but a world nonetheless. In a world in thrall to technological thinking, freedom’s mode of abiding consists for the most part in its withdrawal and quiescence. A manifestation of human being-in-the-world, technological thinking stands in the sharpest possible contrast to what we will now call freedom for responsibility. The latter is also a manifestation of human being-in-the-world, but unlike technological thinking it maintains a certain critical distance between itself and its world. In it, freedom awakes. Technological thinking falls into its world wholeheartedly, becoming its world to such a degree that it is incapable of imagining any other possibility of existence. In a manner that will become clear later, however, freedom for responsibility always remains on the hither side of its world in the form of freedom’s possibilities and freedom’s responsibility. Modern technology, in the sense of technics, has been “captured” by technological thinking to such a degree that the latter has driven the ultimate end of technology as such into darkness and obscurity. It is high time for freedom to rediscover that end—namely, itself—and in so doing to transform modern technology’s essence, its mode of being.

---

102. See ADORNO, supra note 24, at 133.
The End of Technology

We will attempt to show in the next section that there can be only one ultimate end for technology that is worthy of the name: freedom to take responsibility for what humans-and-nature are becoming together. But it is clear that the ultimate effect of technological thinking in present circumstances is precisely the opposite. Marcuse is right in saying that, for the most part, in modern technological societies “the individual’s performance is motivated, guided and measured by standards external to him, standards pertaining to predetermined tasks and functions,” and that “his liberty is confined to the selection of the most adequate means for reaching a goal which he did not set.” As this passage suggests, freedom should never be confused with liberty. If license is the power to satisfy instinctual drives whenever external constraints are absent, liberty is merely the unimpeded power to follow your motives around, like a dog follows its master. But freedom for responsibility has nothing to do with choosing this or that good from a set of goods that has been extracted, usually by someone else, from the standing-reserve that nature and human relations have become. Rather, genuine freedom is openness to destiny—openness to being claimed by something: an aphorism that must remain obscure so long as freedom itself is not thought down to its roots.

IV. THE ESSENCE OF FREEDOM

The phenomenon of freedom shows itself in what the West for two thousand years has called “nature.” Since we are beings-in-nature, the word “nature” necessarily represents the conceptual context in which freedom must be understood, if it is to be understood at all. Thought about human freedom will always remain superficial if it fails to think down to freedom’s ground, which is nature itself. For a long time the idea that man is part of nature has been held captive by a criterion that holds nature to be an intermediate end: nature has been seen merely as the direct means of life and the material object of human activity. This captivity must be broken if freedom is to be understood in a manner that is sufficiently radical. Hence freedom’s grounding question—what is nature?—must be asked in a way that does not treat the answer as self-evident.

Fortunately, a nearly forgotten tradition gives us the means to ask this

103. MARCUSE, supra note 3, at 45.
104. See, e.g., KARL MARX, Economic and Political Manuscripts, in EARLY WRITINGS 63, 126 (T.B. Bottomore trans., 1964).
grounding question. In Western thought, before there ever was a “nature” there was *physis*. This Greek word is the root of our words “physics” and “physical.” Roman thinkers later translated *physis* into *natura*, the antecedent of our words “nature” and “natal.” The translation has been decisive ever since: the image of the natal (Mother Nature giving birth to the world’s many beings) supplanted an altogether different image of self-generation and self-renewal. Lost in the translation was an entire way of thinking about nature. *Physis* was not originally a realm of natural (as opposed to man-made) objects interacting with one another in processes that occupy space and time in a determinate way. Rather, *physis* was conceived as the origin of a kind of unfolding that proceeds, as it were, from out of itself, and that manifests itself in the persistence in being of all that is: a concept that undoubtedly reflects what Werner Jaeger calls the Greeks’ “organic point of view.”\(^{105}\) Moreover, *natura* was destined to become a theological concept in the Middle Ages (as in “God created nature”), whereas the concept of *physis* transcends theology: the Greeks thought that the gods themselves, just like everything else in the world, were manifestations of *physis*, and not the other way around.\(^{106}\)

This rather odd notion of self-generation contradicts our deeply accustomed way of thinking about nature, at least in modernity, as a realm (*natura*) of created beings that rattle around in space and pass through time. *Physis* names *presencing* as the manner in which beings persist in being. Although *physis* itself is radically simple and obvious, it is hard for modern humans to grasp as a concept, thanks to the ubiquity of technological thinking. *Physis* eludes the common sense of technological thinking precisely because it is a condition of common sense: the eye that sees does not, after all, see itself seeing. Consequently, just as there are color-blind people, so too there will be people who are blind to *physis*.\(^{107}\) Nevertheless, we will forge ahead with full confidence in the truth of Spinoza’s epigram: “For all that is

\(^{105}\) Werner Jaeger, *1 Paideia: The Ideals of Greek Culture*, at xx (Gilbert Highet trans., 1945).

\(^{106}\) See Heidegger’s classification of the three stages of Western thought on the being of beings: “[A being] in its emergence unto itself (Classical Greece); [a being] caused by a supreme [being] of the same essence (Middle Ages); [a being as] the extant as object (modernity).” Martin Heidegger, *Contributions to Philosophy (From Enowning)* 120 (Parvis Emad & Kenneth Maly trans., 1999).

The End of Technology

excellent and eminent is as difficult as it is rare.” When the pre-
Socratic philosopher Heraclitus said that “the sun is not only new each
day, but forever continuously new,” he meant to draw attention to
something very uncanny about existence. Foreshadowing Einstein’s
general theory of relativity by two-and-a-half millennia, Heraclitus knew
that the present moment and all of the beings that are present in it
(ourselves included) are not like a uniformly aging box full of its
contents; he saw that they belong together in an inseparable and dynamic
jumble made up of both space and time. From the standpoint of physis,
life and world are never two things standing side by side. Rather, life
plus world equals history: not history in the dry sense of something that
is past, but history in the active sense of the ongoing making of history
in the here-and-now.

If natura’s best metaphor is that of birth, then perhaps the most
appropriate metaphor for physis is that of bud and flower. The
“fecundity of nature” (natura) thus stands opposed to a nature that is like
a rose bud unfolding itself into a blossom (physis). The true salience of
this difference is not to be found in the observation that natura connotes
the female, and physis the male, but in the insight that the image of
fecundity represents a kind of splitting in two, whereas that of blooming
represents the persistence in being of a unity. Moreover, the law of cause
and effect governs natura, while in physis there is something about
nature that precedes all talk of causation: for the ancients, physis is that
which is always already here. If natura gives us the idea of the real (real
beings, real events), then physis gives us the idea of reality as a whole,
which is a lot “bigger” phenomenon than even the biggest real thing.
Reality as a whole manages to persist without its persistence being
caused by any particular real being or event. Although science can hope
to explain the behavior of all that is real, it has and can have nothing to
say about the mind-boggling peculiarity that there is a reality in the first
place.

Continuing with the metaphor of blooming, the indifference of physis
to causation is like that of the rose in this remarkable couplet from The
Cherubic Wanderer, a book of poetry written by Angelus Silesius in the
sixteenth century:

108. “Nam omnia praecelara tam difficilia quam rara sunt,” quoted in ARTHUR SCHOPENHAUER,
109. HERACLITUS, supra note 58, at 13 (Fragment 6).
110. See MARTIN HEIDEGGER, AN INTRODUCTION TO METAPHYSICS 14 (Ralph Manheim trans.,
1959).
The rose is without why: it blooms because it blooms,  
It pays no attention to itself, asks not whether it is seen.111

To be “without why” is not an occult or mysterious state; it means, first and foremost, to persist in being before becoming a mere object for human contemplation or manipulation. Spinoza had a term for this extraordinary tendency of beings to persist in being: conatus essendi.112

Think of this conatus essendi as such and in its own right, and the idea of physis can begin to emerge from the mists of metaphysics into the sunlight of the here-and-now. In any case: for the Greeks, nature as physis must first persist in being before nature as natura can exhibit its dazzling array of causal relations to the inquisitive human mind.

In nature conceived as natura human beings stand above, and ultimately against, non-human beings in the form of nature’s creatures. In physis, however, the emerging, flourishing, and decaying of all beings (human and non-human) are constantly manifesting nature. They are nature, in fact. In natura humans are the first and best of Mother Nature’s creatures, the “sumnum bonum... of creation.”113 In physis humans share with all beings a non-natal origin that Heraclitus described as an ever-living fire, itself without origin, that is constantly being kindled in measures and being put out in measures.114 Nature gives us the idea that we are different from it; physis gives us the idea that we are it. Something is “unnatural” if it does not fall into place in the world conceived as a collection of beings governed by “natural” laws. But the negation of the word physis does not describe anything unnatural in this sense. From a Greek point of view, physis is not so much the antithesis of nomos (lawfulness): rather, physis makes nomos possible by being here in the first place, so that beings can then be observed, by humans, to behave or violate natural or man-made laws.

Modern science and technology are completely, if not obsessively, grounded in the idea of nature as natura. Science is the development of a certain way of looking at things: a way that the Greeks originally called theoria. However, theoretical knowledge for the Greeks (sophia) was

111. The poem is quoted, and information about the poet is provided, in Martin Heidegger, THE PRINCIPLE OF REASON 35 (Reginald Lilly trans., 1991) (1957).


114. Heraclitus, supra note 58, at 25 (Fragment 30).
The End of Technology

sharply distinguished from the knowledge that is appropriate to technē. The latter kind of knowledge, which they called phronesis, is for the sake of production only, whereas sophia lies beyond the sphere of mere utility in the form of seeing and knowing for its own sake.115 Things are otherwise with technological thinking in our era, which demands scientific results (and pays for them, by way of research funding) that can be transformed into a standing reserve of useful things and procedures. In conducting modern science and technology, we are not ourselves: we ask questions the way “everyone” asks them, and we conduct research the way everyone else does.116 From the standpoint of technological thinking, the proud motto of pure science that is carved in the tombstone of the great mathematician David Hilbert—“Wir müssen wissen. Wir werden wissen.” (“We must know. We will know.”)117—leads naturally and inexorably to another, more sinister, motto: “We must control and administer what we know. We will control and administer it.” The scientific conquest of nature thus slides easily into the “scientific conquest of man.”118

A natura that is fully known implies the entry of all natura’s beings into spheres of domination and control.119 This is not so from the standpoint of physis, where things are never first of all objects of theoretical contemplation or political and economic manipulation; in emerging with us into the world, beings are always already here for our dealings with them. As modes of knowing, science and technology accomplish something other than themselves, whereas physis always accomplishes just itself. And although scientific theory in service to technological thinking would like to convince us otherwise, it is only one of the many possible (and valid) ways that we can choose to view and speak about nature in the sense of physis.

Take the example of images produced by the human mind. The psychoanalyst approaches human images scientifically, as symptoms of a cause. Freud’s pedantic reduction of the mystic’s “oceanic feeling” to

---

115. HEIDEGGER, supra note 61, at 84.
118. MARCUSE, supra note 43, at xiv.
119. Cf. HORKHEIMER & ADORNO, supra note 7, at 95 (“Technical rationality today is the rationality of domination.”).
certain experiences and feelings in infancy is a good example of this.\textsuperscript{120} The psychoanalyst and psychologist seek to explain things, and hence they understandably abandon “ontological investigation of the image, to dig into the past of man,” as Gaston Bachelard puts it.\textsuperscript{121} In doing so, however, they more or less inadvertently “explain the flower by the fertilizer,”\textsuperscript{122} for the image itself then loses all meaning apart from its place in the causal nexus. In Heideggerian terms, the sciences of the human mind are characterized by a “forehaving” that interprets being as a series of causes and effects, and by a “foreconception” that addresses and speaks about mental life only in terms of this series.\textsuperscript{123} Both the forehaving and the foreconception are firmly locked into various academic “disciplines” that suffocate all other kinds of questioning. But of course there are, nevertheless, many other possible ways of looking at images. The phenomenologist, for example, aspires to “take the image in its being”\textsuperscript{124} by merely describing what is seen, rather than explaining or analyzing it.\textsuperscript{125} And there are at least some artists, poets, and aesthetes who are content to make or contemplate the image as such,\textsuperscript{126} all the while remaining indifferent to its proper place in the causal nexus, and sometimes even to its potential to turn a profit. In the present age, however, these and other non-technological ways of looking at things remain as but small islands within a vast sea of technological thinking. For humans everywhere are submerged in a technological anthropology that already knows what man is, and hence can never ask the question of whom he might be.\textsuperscript{127}

In nature as \textit{natura} human needs and wants, including the scientific desire to know, are set against a nature that is an adversary to be conquered and tamed. To name nature’s children, to unlock her secrets, to marshal her wealth and resources, to master her processes and make

\begin{itemize}
  \item \textsuperscript{120} \textsc{Sigmund Freud}, \textit{Civilization and Its Discontents} 11–20 (James Strachey ed. \& trans, 1962) (1961).
  \item \textsuperscript{121} \textsc{Gaston Bachelard}, \textit{The Poetics of Space}. at xxvi (Maria Jolas trans., 1969).
  \item \textsuperscript{122} \textit{Id}.
  \item \textsuperscript{123} Martin Heidegger, \textit{Phenomenological Interpretations in Connection with Aristotle}, in \textsc{Heidegger}, \textit{supra} note 113, at 111, 121.
  \item \textsuperscript{124} \textsc{Bachelard}, \textit{supra} note 121, at xxv.
  \item \textsuperscript{125} See \textsc{M. Merleau-Ponty}, \textit{Phenomenology of Perception}, at viii (Colin Smith trans., 1962).
  \item \textsuperscript{126} Cf. \textsc{Wallace Stevens}, \textit{The Irrational Element in Poetry}, in \textsc{Wallace Stevens, Collected Poetry and Prose} 781–92 (1997).
  \item \textsuperscript{127} Martin Heidegger, \textit{The Age of the World Picture}, in \textsc{Heidegger, supra} note 6, at 115, 153.
\end{itemize}
The End of Technology

them do the bidding of humans: this is the essence of our attitude towards nature conceived as *natura*. The concealment of nature as *natura* challenges humans to pry it open and discover its secrets like the oysterman pries an oyster open to find the pearl within. In contrast, the concealment of nature as *physis* is a great mystery and paradox: it is Heraclitus’s “ever-living fire,” that keeps on going and going in a way that would put the Energizer Bunny to shame. If *natura* reports itself only in a form that can be ordered within a system of “information,” *physis* refrains from reporting itself at all: it simply endures. If, in *natura*, truth is the “correct” correspondence between a proposition and the facts, in *physis* the truth of presencing shows itself long before any proposition is uttered, or fact determined.

Although our previous survey of the effects of technological thinking may sound pretty grim, to loosen *natura*’s grip on our minds and to begin thinking of nature in terms of *physis* does not have to transform us into Luddites or revolutionaries. No reasonable person could doubt that human beings need natural resources in order to live and prosper. Nature in this sense is the irreplaceable source of all our food, clothing, shelter, medicine, and art. However cruelly and unequally distributed they are at the present time, scientific discoveries, the fruits of modern technology, and technological planning are obviously essential means for maintaining and enhancing conditions of life on this planet. It is not the technological base that presents the problem: it is the superstructure of technological thinking that we must break with. Like Rousseau,128 we do not urge humanity as a whole to renounce technology’s advances in order to avoid the vices of technological thinking. For it is not a question of either/or here. The most pressing question for humanity is whether we also need, or even are able to recognize, nature in the sense of *physis*. Treating nature as *natura* enables life to live. But treating it (including ourselves) as *physis* can begin to make life *worth* living.

Let’s be blunt: the life of a puppet—even a healthy and well-outfitted puppet—is a degraded way of living. In *physis*, however, lives a spontaneity that puppets cannot know. *Physis alone implies the idea of freedom: it reflects the image of ourselves as the uncaused origin of our future*. Freedom and responsibility are impossible to imagine if we conceive of nature as a box containing entities (including ourselves) that obey immutable laws. Thus, the ancient debate between those who think

---

we have free will and those who think that all of our actions are predetermined keeps on sputtering inside a concept of nature whose essential ambiguity has never been worked out.

As Schopenhauer demonstrates, in a nature conceived as *natura* there is no genuine choice, only its illusion: although there exist motives that can cause actions, the motives themselves exist only if they have been caused by something else.\(^{129}\) This is why freedom cannot properly be classified as the phenomenon of an autonomous and reasonable will,\(^ {130}\) for in freedom of choice (*liberum arbitrium*) the will is always held hostage to a prior judgment between two or more given things according to a predetermined motive which, as Arendt puts it, “has only to be argued to start its operation.”\(^ {131}\) Nevertheless, in the Western metaphysical tradition freedom is usually classified as a kind of uncaused causality; there is causality in accordance with the laws of nature, and then there is “another causality through freedom,” as Kant describes it in the *Third Antinomy*.\(^ {132}\) Although this way of thinking about freedom gives us the image of what Kant calls “an absolute causal spontaneity beginning from itself,”\(^ {133}\) this account of a cause-and-effect relation between freedom and its output logically resembles a squared circle, and is a far cry from the concept of freedom that we are attempting to show here, on the basis of the idea of *physis*.

The genuine freedom of which we speak is not properly classified as a kind of causality: on the contrary, as Heidegger puts it, the real relation is vice versa: “causality is a problem of freedom.”\(^ {134}\) This is because causality is the fundamental category of beings conceived ontically as objects that are simply present, and that then change or perish according to natural laws. Whereas freedom in our sense is a self-blossoming origin of a human world that only *then* passes into the category of causation. An origin in the Greek sense (*archē*) is not a cause.\(^ {135}\) It is not

---


130. Schopenhauer pessimistically made the illusion of freedom, including the feeling that we choose our motives, into the phenomenon of a will that is essentially a blind form of striving, without any ultimate aim. SCHOPENHAUER, *supra* note 108, at 308.


133. *Id*.


135. HEIDEGGER, *supra* note 8, at 85.
The End of Technology

a thing from which some other thing proceeds, but is rather a kind of proceeding in which the origin is constantly passing into the originated. Although Kant himself glimpsed something like this order of priority in classifying causation as one of the categories of our understanding, rather than as a property of things in themselves, he failed to see that causality is grounded in freedom, and not the other way around. Freedom stands on the hither side of the actual in the form of possibilities. Since the past has already devoured its possibilities, the possibilities of which we speak should not be conceived as being outlined beforehand, as if freedom were limited to bringing about one of them. Rather, we must succeed, as Bergson says, in conceiving of freedom’s possibilities in terms of the “radically new and unforeseeable.” Only if we understand freedom in this sense can we break decisively with the view that the present state of affairs in technology is inevitable.

Politics traditionally cares about freedom in its “outer” dimension only. Its context is the decision whether and how to restrain the individual’s freedom of action (negative freedom), or to endow the individual with certain freedom-enhancing entitlements (positive freedom). Standing in contrast to these kinds of outer freedom is the traditional metaphysical concept of what Kant calls “inner freedom.” Just as the feeling of inner freedom can be absent despite the removal of all external restraints, so too the inward experience of freedom that the Stoics celebrated can show itself despite all outer manifestations of freedom’s absence. As Arendt correctly says, however, this traditional conception of inner freedom presupposes “a retreat from the world, where freedom was denied, into an inwardness to which no other has access.” Inner freedom thus conceived is a lonely and ineffectual thing. In contrast, our notion of genuine freedom, freedom for responsibility, radically inverts the traditional understanding of these “inner” and “outer” dimensions of freedom. The shape and condition of the public world is now felt to be inner freedom’s burden and responsibility, while a formerly calm inner freedom feels the need to

136. KANT, supra note 132, at 262–64.
137. HEIDEGGER, supra note 134, at 205.
139. IMMANUEL KANT, Critique of Practical Reason, in PRACTICAL PHILOSOPHY 137, 268 (Mary J. Gregor ed. & trans., 1996).
140. See, e.g., EPICETERUS, A MANUAL FOR LIVING 3–5 (Sharon Lebell trans., 1994).
141. ARENDT, supra note 83, at 146.
pass from its refuge into the outer world of politics. Freedom for responsibility *intervenes*, and having intervened, it feels a tinge of guilt no matter what the outcome is. This is not a paradox but an existential imperative for anyone who takes the idea of responsibility seriously.

The traditional Western definition of liberty as freedom *from* politics (political freedom) enshrines the private sphere as the site where freedom transpires.142 And the solemn guarantee (if not the actual achievement) of fundamental human rights, including the right to privacy, is undoubtedly a genuine advance over tyranny and totalitarianism. But it must be noted that freedom for responsibility is never something that ends in the private sphere, beyond the reach of politics. On the contrary, it *begins* there as the individual’s recognition that the terms “private” and “public” are inadequate categories for grasping the nature and the scope of the responsibility for the world that freedom makes.

True freedom is experienced only in acting, not in willing; it is not experienced as sovereignty over something or someone, but as a spontaneous participation in destiny. Freedom thus conceived is not obedience to Kant’s “moral law within,”143 for this kind of freedom *limits* responsibility to compliance with the moral law. It lets you off the hook if you think you have done “the right thing.” Whereas, if each moment is full of possibilities for freedom that the next moment eclipses, this means that we are destined to live a life suffused with responsibility for *all* of the roads not taken. Freedom’s movement is thus tragic, for as the philosopher Emmanuel Levinas remarks: “The tragic does not come from a conflict between freedom and destiny, but from the turning of freedom into destiny, from responsibility.”144 Life lived in the consciousness of the tragic in this sense is always an occasion for guilt; in fact, freedom and guilt entail one another. In one of his earliest texts, Heidegger helpfully elucidates this point in terms of the phenomenon of conscience, which so often refuses to be soothed by reason’s whisper, “I did the right thing”:

To choose one’s conscience means at the same time, however, to *become guilty*. As Goethe once said, “He who acts is always without conscience, irresponsible.” Every action is at the same time something marked by guilt. For the possibilities of action

142. *Id.* at 149.
143. *Kant*, supra note 139, at 269.
The End of Technology

are limited in comparison with the demands of conscience, so that every action that is successfully carried out produces conflicts. To choose self-responsibility, then, is to become guilty in an absolute sense. Insofar as I am at all, I become guilty whenever I act in any sense.\textsuperscript{145}

In the terminology of Max Weber, to be free is be aware of the ultimate meaning of human conduct, and this means to live in the “knowledge of tragedy with which all action, but especially political action, is truly interwoven.”\textsuperscript{146} Unfortunately, the hackneyed old sports saying, “No pain, no gain,” is also an accurate description of all historical change for the better. Only what Benjamin calls a “quite childish anarchism” could refuse to acknowledge any constraint on the individual; only such a one could fail to see that the very meaning of morality and history becomes impossible to think if action, including the violence of radical change, is removed from the sphere of thought and politics.\textsuperscript{147} Conversely, only an ideologue utterly and blindly convinced of the rightness of his cause could fail to see the pain and disruption that his effective and even “reasonable” actions cause. As Levinas puts it: “There are cruelties which are terrible because they proceed from the necessity of the reasonable Order. There are, if you like, the tears that a civil servant cannot see.”\textsuperscript{148} Freedom for responsibility is thus a way of living that is radically different from mindlessly falling into the everyday world of production and consumption, where everyone does pretty much what everyone else expects them to do, and nature becomes a mere assembly of objects for us to use and abuse, conserve and consume, honor and dishonor, all according to the preferences that technological thinking allows and encourages us to form. To transcend this world requires both the courageous action of a hero and the humility and sorrow of a saint.

Although we obviously need nature’s resources, we do not need technological instruments \textit{as such} nearly as much as we need freedom for responsibility. While the attempt to fly from freedom may quench the fires of responsibility and guilt, at least for a while, such an attempt is none other than a cowardly acquiescence in the status quo. As long as

\begin{footnotesize}
\begin{enumerate}
\item Heidegger, \textit{supra} note 116, at 169.
\item BENJAMIN, \textit{supra} note 101, at 284.
\item EMMANUEL LEVINAS, \textit{BASIC PHILOSOPHICAL WRITINGS} 23 (Adriaan T. Peperzak et al. eds., 1996).
\end{enumerate}
\end{footnotesize}
we think of nature exclusively in terms of *natura*, neither humans nor nature will ever be free. Just as the institution of slavery demeans both slave and master, so too treating nature merely as *natura* demeans both nature’s creatures and the human beings who exploit or protect them. From the standpoint of *natura*, environmentalism will never be more than an emotionally driven “preference” to protect and preserve some of Mother Nature’s beings while ignoring or decimating others. According to the scientific maxim “What matters gets measured,” the earth has to get completely mapped and measured before the reality of environmental degradation can acquire any political currency.\(^{149}\) Nature recedes, as it were, so that her beings may be computed and counted in that simulacrum of her that is called “the environment.”\(^ {150}\) But from the standpoint of *physis*, the environmentalist, no less than the anti-environmentalist, freely decides to take a stand within existence, and steers it, for good or for ill, in a direction. If *natura* gives humans the power to dominate nature, *physis* gives nature and humanity the warrant, and the responsibility, to be what they are becoming together.

Let’s be clear: We do not advocate an attitude towards nature that makes it into a “utopian realm of hope” that would miraculously rescue people from the coldness of the market, and from an increasingly frenetic home life that is being “devoured by the monster of competition.”\(^ {151}\) Nor do we idolize nature the way certain fascist writers did, as a mystical place where blood and race are linked to soil, and where sentimentality for one’s “natural” land leads to brutality in defending and extending it.\(^^{152}\) The longing to fly to nature, and to drown one’s individuality and responsibilities in it by becoming “one with nature,” has nothing to do with our concept of freedom for responsibility. We say, on the contrary, that an awareness of nature as *physis* can bring human beings genuinely back to their home and work, back to their individuality, and, more importantly, back to a sense of responsibility. Being attuned to nature as *physis* does not automatically make us free and responsible to any particular being or segment of nature. Rather, it makes us responsible for the nature that we ourselves


\(^{150}\) See Heidegger, supra note 26, at 244 (commenting on the “unfathomable, essential difference between the relationship to a ‘world’ [Weltbezug] and to an ‘environment’ [Umgebungsbezug]”).

\(^{151}\) Leo Lowenthal, *Knut Hamsun*, in *The Essential Frankfurt School Reader*, supra note 65, at 319, 322 (attributing this attitude to Ibsen).

\(^{152}\) Id. (criticizing the work of the proto-fascist Scandinavian writer Knut Hamsun).
The End of Technology

manifest by our actions and inactions. We are like architects whose buildings are fated to be lived in by everyone and everything. The thought of nature as physis yields the chance to see oneself and one’s choices as mattering to history. While this chance does not necessarily lead to tree-hugging or a sense of stewardship over nature, neither does it necessarily lead to their opposites. And that is precisely why we need the concept of nature as physis: to reawaken the monumental insight that we have, indeed that we are, our possibilities.

The freedom of which we are speaking transcends technological thinking to participate as a creative actor in an ever-blossoming world that no single person ever “selects,” but for which everyone is, paradoxically, responsible. Choosing freedom thus means saying “Yes!” to the manner in which you-and-nature continue to persist in being. It means first of all having the time and the inclination to think, and form an idea of, the ultimate end of technology. It also means holding the unity that is the world-and-you accountable for having ignored the question of ultimate ends for so very, very long. Although we risk being mocked in repeating Marcuse’s words, in an object world that is at once humanized and naturalized in this way one can begin to sense “that an electric garbage disposal unit has a ‘soul,’ that there can be tenderness in an automobile, and that a bulldozer can not only pulverize but also restore nature.” In freedom for responsibility, we forget that everything is caused, and we make decisions that destiny fulfills as freedom’s inescapable counterpart. Only the one who succumbs to the belief that technological thinking is inevitable surrenders to fate. But destiny is beyond fate. As Martin Buber puts it, “destiny is not a dome pressed tightly down on the world of men,” but rather freedom’s very meaning and fulfillment. Human beings who are free for responsibility believe that destiny needs them rather than controls them.

Genuine freedom of this sort is hard to glimpse and maintain in the best of circumstances, though, and it is especially hard for those whose daily struggle for existence is more than just an abstraction written about in an academic paper. We must now confront the relation between what Roosevelt called “freedom from want” and the ultimate end of

153. MARCUSE, supra note 2, at 133.
155. See Franklin D. Roosevelt, The Four Freedoms Speech, 1941, in 2 THE PEOPLE SHALL JUDGE: READINGS ON THE FORMATION OF AMERICAN POLICY 751, 756 (The Staff, Social Sciences, the College of the University of Chicago eds., 1949).
technology that we have identified, freedom for responsibility. Although these two freedoms are not the same, we will claim that they are bound to one another inextricably.

V. THE RELATION BETWEEN FREEDOM FOR RESPONSIBILITY AND FREEDOM FROM NECESSITY

Although there is probably nothing that “bores the ordinary man more than the cosmos,” as Walter Benjamin once remarked, it nevertheless is true that we are cosmologically responsible for the shape of our world whether or not we are aware of it, or bored by it. And it is also true that there is no logically necessary connection between the consciousness that one is free for responsibility (as we have described it) and whether one’s material life circumstances place one comfortably beyond the struggle for existence. Experience teaches that even a poor person or a miserable slave can become a philosopher. Nevertheless, neither the threat nor the reality of an empty belly correlates well with a sense of responsibility that extends beyond finding the next meal. Although the Stoic’s advice always to identify the real with the desirable may bring occasional spells of tranquility to the souls of the unhappy and the downtrodden, it is also a recipe for resignation to the effects of technological thinking. From the standpoint of genuine freedom, acceptance of the status quo is the same as a guilty complicity in its perpetuation.

Before people can make a mature decision about the ultimate value of the world that technological thinking has made, they must have both time for reflection and the inclination to reflect. The former is determined by necessity, whether real or artificially perpetuated—a point that we will return to shortly. But the latter requires the mind to make a decisive break with the self-evident naturalness of the world. Only such a break gives thought the critical distance that is necessary to think in terms of technology’s ultimate end. Yet how very hard this is! According to Max Weber’s well-known thesis, the “worldly asceticism,”

156. WALTER BENJAMIN, THE ARCADES PROJECT 102 (Howard Eiland & Kevin McLaughlin trans., 1999).
157. For example, the great stoic philosopher Epictetus was a slave. Moses Hadas, Introduction to ESSENTIAL WORKS OF STOICISM, at vii, xv (Moses Hadas ed., 1961).
158. “Seek not that the things which happen should happen as you wish; but wish the things which happen to be as they are, and you will have a tranquil flow of life.” Epictetus, The Manual, in ESSENTIAL WORKS OF STOICISM, supra note 157, at 85, 87.
The End of Technology

which is characteristic of the earliest stages of capitalist production in Europe, had a religious basis. For the Puritan, Weber writes, “labour came to be considered in itself the end of life.”\(^{159}\) The early, religiously motivated capitalist saw repression of all forms of “spontaneous, impulsive enjoyment” as a moral imperative,\(^{160}\) in part because he thought “[u]nwillingness to work is symptomatic of the lack of grace.”\(^{161}\) The decline and disappearance of this religious basis of early capitalism did not, however, diminish the social importance of work in capitalist production. Nor did it alter the attitude of technological thinking about the relationship between work and worth. As Weber puts it, in two oft-cited passages: “The Puritan wanted to work in a calling; we are forced to,”\(^{162}\) and “the idea of duty in one’s calling prowls about in our lives like the ghost of dead religious beliefs.”\(^{163}\) The individual who does not experience the need to work as a religious duty “generally abandons the attempt to justify it at all,”\(^{164}\) says Weber, thereby confirming our earlier hypothesis that one of the prime effects of modern technology is to channel mental energy away from the consideration and formulation of ultimate ends. People learn that if they think too consistently and too vigorously about their life and where it is going, they can cause themselves all kinds of short-term inconvenience, without any assurance that things can be improved.

Weber’s central insight that reason’s capacity to think about ultimate ends atrophies wherever technological thinking holds sway suggests that a precondition for the recognition and spread of the genuine freedom that we have attempted to uncover—freedom for responsibility—is a material way of life that has been released from the grip of what Marcuse calls “surplus-repression.”\(^{165}\) Work is repressive to the extent that it forecloses the natural and creative ends that individuals would pursue if their time were truly their own. And it ceases to be repressive whenever an increase in the leisure time available to individuals, corresponding to overall increases in the efficiency of socially necessary labor, is interpreted as a fundamental human right and a cause for

\(^{159}\) Weber, supra note 9, at 159.  
\(^{160}\) Id. at 119.  
\(^{161}\) Id. at 159.  
\(^{162}\) Id. at 181.  
\(^{163}\) Id. at 182.  
\(^{164}\) Id.  
\(^{165}\) Marcuse, supra note 2, at 22, 140.
celebration, rather than as an occasion for shame and self-hatred on account of being “unproductive” or out-of-work.

As the overall amount of time required for socially necessary work continues to decrease as a consequence of automation and other technological developments, technological thinking is responsible for allowing the realm of unnecessary repression continually to grow: as if the two were inversely (and perversely) correlated. For despite the fact that the overall efficiency of the system has improved, the system still requires its unemployed members to “earn a living” in some job or other, no matter how degrading and unnecessary it may be when considered from the standpoint of satisfying vital human needs. Enter the McDonaldization of the workforce in the form of low-paying service jobs, the growth of part-time employment without benefits, and the frantic pace of a work ethic that is keyed to the norm of producing short-term profits for shareholders. This surplus repression is irrational from the standpoint of any ultimate end that values humans over things, since increased technological efficiency in the production and distribution of necessary goods ought to lead to more freedom from necessity, rather than the same amount or less.

Technological thinking makes us prove our worth, to others and to ourselves, in the labor market. But why should the labor market decide our worth if a free human life is the ultimate end of socially necessary labor, rather than just a means to the end of production for its own sake? In a recent essay on Marcuse’s thought, Jürgen Habermas refers to a widespread estimate that in the countries comprising the Organization for Economic Cooperation and Development (OECD), the gross national product could theoretically be produced by twenty percent of those able to work. The social meaning of this estimate is obvious from the standpoint of any plausible non-Puritanical conception of technology’s ultimate end. As Habermas says, with notable understatement, “if a constantly increasing part of the working population becomes ‘superfluous’ for the reproduction of society, the close connection between occupational success and social recognition

166. Thirty member-states belong to the OECD: Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, South Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovakia, Spain, Sweden, Switzerland, Turkey, United Kingdom, and United States. See http://www.nationmaster.com/index.php?kp=OECD (last visited Dec. 23, 2003).

167. Jürgen Habermas, The Different Rhythms of Philosophy and Politics for Herbert Marcuse on His 100th Birthday, in MARCUSE, supra note 2, at 233, 236.
The End of Technology

can hardly be maintained.” 168 And even if the estimate on which Habermas relies is off by a lot, it cannot reasonably be denied that technological improvements in the efficiency of production and distribution are tending in the direction he describes. Nor can it be denied that we are doing very little to anticipate and prepare for the effects that this technological cornucopia will (or should) have on the relations between work and worth, rich and poor, and Self and Other. Instead, we seem content to accept an absurd state of affairs in which “the power of the system over human beings increases with every step they take away from the power of nature.” 169

Habermas is speaking primarily about the surplus repression of work in highly developed nations. Looking beyond repressive conditions there, one encounters a similar but far more brutal surplus repression in undeveloped and developing regions. This happens to be a world where a large portion of the human population is ravaged by curable diseases, 170 where eight hundred million people suffer from avoidable hunger and malnutrition, 171 where 1.2 billion people live on less than one dollar per day, 172 and where poor working parents all too often have to make the untenable (and technologically unnecessary) choice between protecting their children from harm, including taking care of them when they are ill, and keeping their jobs. 173 The fact that in such a world technological thinking considers overcapacity in the American automobile industry, currently at twenty percent and growing, to be bad news 174 speaks volumes. This fact should be more of a cause for global

168. Id.
171. Beth Potier, UN’s Diouf Addresses Hunger: “The World Hunger Problem Is Clearly Political, Not Technical,” HARV. U. GAZETTE, Feb. 6, 2003, available at http://www.news.harvard.edu/gazette/2003/02.06/10-hunger.html. The article notes that equal distribution of all the world’s food supply would “provide everyone on the planet with a robust 2,800 calories per day, a 17 percent increase from 30 years ago.” Id.
173. Jody Heyman, Families on the Edge, HARV. MAG., July–Aug. 2003, at 50, 53 (reporting on a global study of this problem and concluding as follows: “The world can afford to ensure adequate working conditions and social supports for families; what we can’t afford is neglect.”).
celebration than it is a cause of short-term worry, since it indicates that
technology has at its command an enormous reservoir of untapped
potential that could be used to liberate the universal community of men,
women, and children from the oppression of an unnecessary struggle for
existence. Where, except in a madhouse, could advertisements for
$3,000 watches and $17,000 pearls sit side-by-side with a news report
about the cash crisis in global AIDS funding,\textsuperscript{175} without evoking the
profoundest sense of shame and outrage? Where, except in a world that
has been blinded to ultimate ends by technological thinking, could the
observation that the annual output of Liberia “is roughly what
Americans spend on skiing equipment”\textsuperscript{176} fail to elicit feelings of guilt
and disgust? When we in the prosperous West think of AIDS in Africa
and the millions we allow to die there, the ugly suspicion arises that this
“letting die” is, as Jacques Derrida says, a form of killing.\textsuperscript{177} The
suspicion arises that, in an era of preventable death and disease, “Thou
shalt not kill” means that thou shalt cause these others to live.

If the Enlightenment watered the acorn that became the oak of
technological thinking (the essence of modern technology), then it also
gave us the profound idea and aspiration of \textit{universal} human
emancipation. In Kant’s universal ethical subject, bound to be oth-
regarding by reason in the form of the Categorical Imperative,\textsuperscript{178} in
Hegel’s idea that the master-slave relation is destined to be dialectically
overcome (\textit{Aufgehoben}) by a higher stage of development in which
universal human recognition prevails,\textsuperscript{179} in Montesquieu’s correlation of
the \textit{ability} to be free with true freedom of will,\textsuperscript{180} and in Marx’s
description of freedom as the development of human potentiality for its
own sake, beyond the realm of necessity,\textsuperscript{181} one can find the origins of
the tattered but still lofty progressive notion that technology should aim
at \textit{universal human emancipation from necessity}. The universality of this

\begin{itemize}
  \item \textsuperscript{175.} See N.Y. TIMES, July 17, 2003, at A2–A4.
  \item \textsuperscript{176.} \textit{Why Liberia Is Not Somalia}, THE ECONOMIST, July 19, 2003, at 11.
  \item \textsuperscript{177.} GIOVANNA BORRADORI, PHILOSOPHY IN A TIME OF TERROR: DIALOGUES WITH JÜRGEN
  \item \textsuperscript{178.} KANT, supra note 101, at 195.
  \item \textsuperscript{179.} G.W.F. HEGEL, PHENOMENOLOGY OF SPIRIT 111–19 (A.V. Miller trans., 1977); see also
  ALEXANDRE KOEVE, INTRODUCTION TO THE READING OF HEGEL 43–44 (James H. Nichols, Jr.,
  \item \textsuperscript{180.} “\textit{La liberté ne peut consister qu’à pouvoir faire ce que l’on doit vouloir}” (“Liberty consists
  in being able to do what one ought to will”), quoted in ARENDT, supra note 83, at 161.
  \item \textsuperscript{181.} T.B. Bottomore, \textit{Introduction} to MARX, supra note 104, at vii, x.
\end{itemize}
ideal of universal freedom from necessity is both challenged and measured by the way in which the parts of the historical whole relate to those who, as Žižek puts it, lie “at the bottom.” In short, the ideal requires an effort to satisfy vital human needs on a global scale.

The particular historical forms in which the aspiration for universal human freedom has appeared to date—primarily economic liberalism rooted in mass democracy and totalitarian state communism—do not exhaust its potential to inspire us. If economic liberalism finds freedom only in its negative sense (the right to be left alone by the state in the pursuit of market relations), and if state communism finds it in the perverted positive sense of an enforced conformity to social goals, then the ideal of universal freedom from necessity needs to find its idealization and its expression in a way that transcends both discourses. For when instrumental reason separates from the ultimate purpose of thought, the ideal of knowledge through a “method” becomes tautological, “since it does no more than fulfill the demands of method.” The biblical admonition “Seek, and ye shall find” becomes the guideline to seek only what you will find according to this or that “useful” method.

It is unlikely that help in our project will be found in the direction of modern analytic philosophy. Analytic philosophy’s primary tendency is to deflate universals like “freedom,” “universal humanity,” and “technology,” reducing them to descriptions of particular given operations of speech and conduct: Wittgenstein’s language-games, for example. Marcuse calls this mode of philosophizing “operational,” defined as the tendency to make concepts synonymous with a corresponding set of operations. And while this technique may clarify and correctly describe linguistic usages, in the end the equation of language and its function is complicit in perpetuating the dominion of technological thinking. For the function of a thing in a society that is governed by technological thinking is itself technological, a point that is

182. See ŽIŽEK, supra note 40, at 224.
183. ADORNO, supra note 24, at 76.
184. Matthew 7:7 (King James).
186. MARCUSE, supra note 43, at 87.
187. Id. at 203; see also ADORNO, supra note 24, at 76 (“[T]he only productive knowledge is that which goes beyond pure analytical judg[]ment, which transcends this operational-tautological character.”).
missed or glossed over by such thinkers as Ronald Dworkin, who equates the meaning of his philosophical concept of “right answers” in law with the “ordinary” usage of lawyers. 188 In law, operationalism expresses itself in the form of Hohfeld’s idea of “operative facts”: these are never meant to be objective indices of being or reality, but are only the result of a legal operation that adduces exclusively those kinds of facts that the law predetermines as sufficient to alter legal relations. 189 If the philosophical concept of right answers to legal problems is rigidly circumscribed by what everybody in “real life” says it is in this manner, then there is nothing for critical thought to do except throw in the towel and become a lawyer in ultimate service to the given reality. 190

Universals like “freedom” have always had the potential to reproach what is by juxtaposing it with what might be. Consider, for instance, what Martin Luther King, Jr., was able to do with the words “freedom” and “justice.” If legal and philosophical analyses succeed in demythologizing universals, then they will have rendered them incapable of invigorating the imagination and moving people to oppose the given reality. Whether they mean to do so or not, the atrophied kind of philosophizing and the operational kind of lawyering described in the previous paragraph tend to make the emancipating effects of universals harmless to the system. Such thought represses critical reason’s ability to think beyond operational terms by criticizing as “metaphysics,” “disguised ontology,” or “irrelevant” any view that does not accept the existing order of things, or that chooses to opt out of the stubborn adherence to immediate facts. That law apologizes for the existing order in this way is hardly surprising. But when philosophy makes the same kind of operational moves that law does, it too offers a kind of apology for the status quo—contrary to the proud idea, originating in Socrates, that the philosopher’s proper role is to act as a gadfly to society and the state. One feels compelled to wonder (or lament) where critical reason has gone, until one remembers that even philosophy can succumb to technological thinking. Beyond all language analysis of this sort, beyond

190. For more on this point, see Louis Wolcher, Ronald Dworkin’s Right Answers Thesis Through the Lens of Wittgenstein, 29 RUTGERS L.J. 43 (1997).
law, and beyond all political “third ways,” we will argue that the idea
of universal freedom from necessity should be seen as genuinely
mythological, and not merely utopian, metaphysical, or juridical.

The distinction between myth and utopia was first drawn in
Reflections on Violence, which was written at the beginning of the
twentieth century by Georges Sorel, a French anarchist philosopher and
former civil engineer for the French state. J.B. Priestley once remarked
that if one could only grasp why a retired civil servant wrote such a
book, then the modern age could be understood. In any case, in
Reflections on Violence, Sorel identifies the mythological with the
language and imagery of any individual or social movement that is
convinced things must change, but is not sure how it can be changed. In
contrast, he calls utopian any program of reform that lays down in
advance the exact shape of the world that is to be achieved. Sorel
sharpens this critical distinction in the following passage:

A myth cannot be refuted since it is, at bottom, identical to the
convictions of a group, being the expression of these convictions
in the language of movement; and it is, in consequence, unanalysable into parts which could be placed on the plane of
historical descriptions. A utopia, on the other hand, can be
discussed like any other social constitution; the spontaneous
movements it presupposes can be compared with those actually
observed in the course of history, and we can in this way
evaluate their verisimilitude; it is possible to refute it by
showing that the economic system on which it has been made to
rest is incompatible with the necessary conditions of modern
production.

When mythological thinking becomes utopian it exercises its freedom of
action in a direction. Prior to that, it dwells at a stage where anything is
possible, and nothing is decided yet.

As Sorel’s text implies, technological thinking can always discredit

191. The term “third way” refers to the political slogans of those who, like Britain’s Tony Blair,
seek to get the votes of disillusioned liberals and middle-of-the-road conservatives by promising the
conservatives capitalism with a compassionate heart, and promising the liberals socialism infused
with “rationality” and market discipline. See Bagehot: A Poodle That Snaps, THE ECONOMIST, July
19, 2003, at 46.
193. Id. (frontispiece).
194. Id. at 27–31.
195. Id. at 29.
utopias on the ground that they are unrealized or unrealizable ends, given the present factual starting point. But its attempt to disparage myths, in Sorel’s sense, is always fated to be ineffectual. A myth cannot be refuted so long as it represents the deepest convictions of those who reject the present state of affairs as beneath the dignity of man.  

Technological thinking and mythological thinking both agree that the realm of facts (the given reality) stands opposed to the realm of myth (what is not). But technological thinking construes the rational implication of this opposition to be acquiescence in the status quo, whereas mythological thinking finds its motives, and its hope, in what is not yet a reality. Although it may sound like a shocking exaggeration to say it, only in this way can mythological thought, afloat in a sea of technological thinking, avoid agreeing with Camus that the last genuine philosophical problem remaining today is whether to commit suicide.

While Sorel was writing specifically about the conflict, in fin-de-siècle Europe, between parliamentary reformism (some of which was based on utopias) and the revolutionary social movements of his day (many of which were based on myths), his distinction between myth and utopia transcends that particular historical context. It gives us a useful conception of the meaning of responsibility in conditions that appear impervious to change. For the poverty of any particular utopian vision is nothing compared to the mythological conviction that the end of technology is universal human emancipation. And with such a conviction comes responsibility for the way things are no matter how difficult or impossible movement in the direction of the myth may now appear to be. Those in the grip of a fanatical utopian vision are capable of the most unspeakable outrages, as Sorel correctly observes. This observation implies the large grain of truth in Nietzsche’s aphorism that “the actual man represents a much higher value than the ‘desirable’ man of any ideal hitherto.” Nevertheless, to recognize a particular utopia as an aim is never a matter of freedom, but of right or wrong judgment, just

196. Id. at 30.

197. ALBERT CAMUS, The Myth of Sisyphus, in THE MYTH OF SISYPHUS AND OTHER ESSAYS 3, 3 (Justin O’Brien trans., 1955) (“There is but one truly serious philosophical problem, and that is suicide.”).

198. SOREL, supra note 192, at 10 (“During the Terror [of the French Revolution] the men who spilt the most blood were precisely those who had the strongest desire to let their equals enjoy the golden age of which they dreamt and who had the greatest sympathy for human misery.”).

as the power to command is also not a function of freedom, but only of strength or weakness. 200 Whereas those who are moved by the myth of universal human emancipation freely accept guilt for the consequences of what is done and undone in the name of this myth. Like Arendt’s “principle” of freedom,201 the mythological does not come from within, but inspires from without. If the one who believes in universal human emancipation moves now slowly and incrementally, now rapidly and radically, such a one always does so in the form of a freedom that knows itself to be both a “beginning and a beginner,”202 and that knows it alone is responsible for the world that freedom makes. Such a one knows that without the violence of change, there can be no progress, and that without the compassion that comes with a sense of responsibility, there can be no progress that is worthy of the name. If the only thing worse that the nonexistence of a just existence for humanity is the nonexistence of what Benjamin calls “the irreducible, total condition that is ‘man,’”203 this can only mean that freedom for responsibility holds both justice and life in a tragic balance.

The myth (in Sorel’s sense) that the awesome power of modern technology can be used to liberate universal humanity from the coils of necessity thus connects the notion of freedom for responsibility with a certain kind of ethical responsibility for the Other. We do not refer to traditional conceptions of ethics based on norms or consequences, but rather to the radical kind of ethical responsibility that is described in the work of Emmanuel Levinas. Freedom for responsibility, as we have portrayed it up until now, is existential freedom: it is linked to the manifestations of the world (being) before it is linked to any of our fellow human beings (the Other). But as is more fully explained elsewhere,204 Levinas places ethical responsibility for the Other ahead of any other existential comportment—indeed, he places it in a logical category that he calls, paradoxically, “otherwise than being.”205 While this is not the place for an extended discussion of the extraordinary subtlety and originality of Levinas’s thought, it is appropriate to say that his central phenomenological insight is that we feel ethically responsible

200. ARENDT, supra note 83, at 152.
201. Id.
202. Id. at 170.
203. BENJAMIN, supra note 101, at 299.
205. LEVINAS, supra note 148, at 114–15.
for our fellow human beings before any explicit calculation based on legal or moral norms. Indeed, our practice of making explicit, norm-based, calculations of whether or not responsibility is owed is one of the principal ways we avoid or shed this primordial sense of responsibility.206 Contrary to an entire tradition of political philosophizing that originated with Hobbes,207 Levinas holds that human beings always experience themselves as sociable before they experience themselves as free.208

One does not have to rank existential freedom and ethical responsibility for the Other in terms of priority, however, in order to see that at the very least the former tends towards the latter, and vice versa. If I live with a sense of responsibility for the shape of the world, then the avoidable suffering of others must always show itself as one of the consequences of my actions and inactions; thus, only if I see myself as a monster or a swaggering and egoistic Übermensch will I fail to be moved by their plight. Likewise, the one whose heart always bleeds for the suffering of others has the strongest of incentives to exercise his existential freedom by attempting to change any state of affairs that is responsible for unnecessary suffering.

The connection between freedom and ethical responsibility that has just been drawn is not offered as a necessary implication. Unlike John Rawls, we do not seek to deduce from a self-regarding conception of reason, operating from behind the veil of ignorance, the logical conclusion that our institutions must be set up in such a way as to improve the condition of the worst off in society.209 Nietzsche, for one, would rightly object that the more-than-ample historical record of human brutality and greed demonstrates that freedom and practical reason, in the metaphysical sense of will-to-power, tend to negate the necessity of any such conclusion.210 Nietzsche to the contrary notwithstanding, however, not only do we maintain that the connection we have drawn is plausible “empirically,” but also we offer it as deeply mythological in Sorel’s sense.

208. LEVINAS, supra note 148, at 91 (asserting “the fact of human fellowship, prior to freedom”).
209. See RAWLS, supra note 88, at 150–61 (arguing for the “veil of ignorance” and “maximin” rules).
210. See NIETZSCHE, supra note 199, passim.
As for the empirical, we submit as evidence the countless large and small kindnesses that daily pass between strangers, before being reduced to “history” and “systems of ethics.” We also offer the ubiquitous phenomena of compassion and conscience in the face of the curable suffering of others. True compassion is not an experience that transpires between this person and that person conceived as an I and an It, for the word “experience” implies the secondary status of the things that are related through experience. Like Martin Buber’s concept of I-Thou, compassion is not an experience but a primary human relation. It is a connection that is discovered, not created, a bond that is found, not forged. Unfortunately, as Buber correctly observes, the development of the function of experiencing things and using them technologically comes about mostly through a decrease in the power to discover and sustain relations of this sort. The historical struggle for existence and modern technological thinking, both quintessential examples of “experiencing and using” in Buber’s sense of the abstract I-It relation, thus tend to drive the primary relation of compassion underground. However, we would contend that the depressing historical record of human selfishness and brutality in our age and prior ages should be seen as ambiguously relevant counter-evidence at best, especially if we are right that modern technological thinking artificially perpetuates the struggle for existence.

But what if we are not right, a critic might ask, as a matter of “fact” about the central importance of sociality and compassion in the way that humans are constructed—what then? More important than any proof that is based on a pessimistic, and ultimately technological, assessment of so-called “human nature,” we say that the expression of the intimate linkage between freedom for responsibility and universal freedom from necessity is, in the final analysis, profoundly mythological in Sorel’s sense. It represents the conviction of anyone who thinks the world has gone mad in the many ways that we have attempted to demonstrate. It authorizes us to say that the meaning of freedom for responsibility requires freedom from necessity in order to flourish, and naturally leads to the feeling that one is responsible for the artificial perpetuation of necessity and unnecessary suffering in the world. It is nothing other than

211. BUBER, supra note 154, at 19–44.
212. Id. at 52.
213. Id. at 54 (“If a man lets [the I-It relation] have the mastery, the continually growing world of It overruns him and robs him of the reality of his own I, till the incubus over him and the ghost within him whisper to one another the confession of their non-salvation.”).
the shape of hope for what might be in the face of despair over what is.

VI. CONCLUSION

Let us end where we began, and ask once again: What is the end of technology? What this question asks in its deepest signification is nothing technological. Although technology in the largest sense of the word consists of human beings uncovering beings and making them accessible, technology as such is not ultimately for the sake of any particular project or outcome. Technology thus conceived exists ultimately for the sake of human freedom, which can only express itself in and by means of technology. Genuine freedom is freedom for responsibility: the kind that knows that it is freedom itself that constructs the very obstacles to freedom that continue to confound a suffering humanity. Freedom from necessity is not a logical condition of the possibility of freedom for responsibility, but it is nonetheless an empirical condition of the latter’s spreading and flourishing. The harried and overworked, the hungry, the sick, and the downtrodden: their gnawing need to survive a struggle for existence that modern technological progress could make unnecessary relentlessly devours their chances to live in freedom. Thus, not only does freedom begin only when “the I-will and the I-can coincide,” as Arendt puts it, but it is also the case that those who are already free for responsibility wear other people’s avoidable lack of freedom as a shameful stain on their sense of responsibility.

Marx once uttered the profound truth that “we have to emancipate ourselves before we can emancipate others.” If this essay has achieved its goal, then its most important argument will have landed on fertile soil: The grip that technological thinking exercises on our imagination must be broken, and technology must return to the status of a means to the ultimate end of universal human emancipation. All progressive thought about the consummation and the ultimate end of technology requires, first and foremost, what Marcuse calls a “consciousness of the blatant contradiction between the scientific-technological possibilities and their destructive and repressive realization.” This consciousness means the ability to think the category of individualism beyond the

214. ARENDT, supra note 83, at 160.
215. KARL MARX, On the Jewish Question, in MARX, supra note 104, at 1, 5.
216. MARCUSE, supra note 2, at 139.
The End of Technology

making of market choices that implement a life of unnecessary toil and servitude. It means defining individuality in terms of being able to speak one’s own language, to have one’s own emotions, and to follow one’s own heart. And it requires the conviction that true individualism can only begin beyond the realm of necessity: a truth that well-paid tenured legal academics ought to know from personal experience. To those who would disparage and scorn this vision of human emancipation as unrealistic and naively utopian, we say: So what? We have seen the world that technological thinking has made, and the irrationality of its rationalism shows itself to the faculty of judgment in a way that is every bit as dysutopic in fact as our idea of universal human emancipation is mythological and “unrealistic.”

Like Kierkegaard’s “knight of faith,” the ones who are free for responsibility in the way we have attempted to describe make a kind of three-fold movement. First concentrating their infinite desire for a world that transcends technological thinking by acting to achieve such a world, they quickly learn the tragedy of unintended consequences when their best-laid plans go awry. This leads them, sadly, to a second movement: resignation to the finite reality that their actions are always bound to produce imperfect, and sometimes even disastrous, results, and that they are responsible for these results. But then occurs “still another movement more wonderful than all,” as Kierkegaard puts it. After the two-fold movement of desire and resignation, the ones who are free for responsibility absurdly accept the truth of their impossible aspiration for universal human emancipation. They do not (necessarily) do this because they have faith in a God who is believed to make all things possible, as Kierkegaard himself would have it, or because they are scientifically convinced in the inevitability of progress towards universal human emancipation, as were Hegel and Marx. Indeed, they know that the source of true freedom remains present, as Arendt puts it, “even when political life has become petrified and political action impotent to interrupt automatic processes.” They accept the truth of their aspiration for universal human emancipation simply by virtue of their irrational faith in the fundamental goodness of humanity itself.

218. Id.
219. Id.
220. ARENDT, supra note 83, at 169.
The empirical possibility that human beings cannot enjoy the continuing benefits of modern technology without suffering the continuing hegemony of technological thinking is simply too horrible to be believed or credited, and this is not just because positivism and empiricism, as tools of technological thinking, tend to conflate the non-actual with the impossible. Cicero’s bold remark to the effect that he would rather go astray with Plato than hold true views with Plato’s opponents captures the essence of our attitude towards the positivist’s “bitter with the sweet” argument. For the terrible idea that technical progress can never be humanized and naturalized in the ways that we have tried to describe in this essay represents the death of all hope, and of all critical thinking. Sadly, the realization of this idea would truly represent a miserable human history that ends not with a bang, but a whimper.

Pursuing no conquest, the social movement to pacify the unnecessary struggle for existence in the developed world, and to satisfy, by means of technology, vital human needs on a worldwide scale needs to make no plans for utilizing its victories. This statement does not reflect an attitude of recklessness and irresponsibility, nor does it abjure the need for careful calculation and planning. Rather, the statement is predicated on the truisms that “universal” means universal, and that the very notion of “victory” implies the antithesis of universality: the scission of humanity into permanent camps of winners and losers. It rejects the egoistical sentiment “after me the deluge,” in favor of a slogan that in its deepest sense transcends the status of an aphorism to become mythological. The slogan’s particular expression (though not its content) was first rendered by Herbert Spencer. It is just this: No one can be perfectly free until all are free.

223. The reference is to T.S. Eliot’s extremely apt poem The Hollow Men, which concludes with the following lines: “This is the way the world ends / This is the way the world ends / This is the way the world ends / Not with a bang but a whimper.” T.S. ELIOT, COLLECTED POEMS 79, 82 (1963).
224. See SOREL, supra note 192, at 161 (using similar language to refer to the proletarian general strike).
225. See id. at 93 (“Our bourgeoisie desire to die in peace—after them the deluge”).
SOME REFLECTIONS ON LONG-TERM LESSONS AND IMPLICATIONS OF THE ACCESS TO JUSTICE TECHNOLOGY BILL OF RIGHTS PROCESS

Richard Zorza*

The Washington State Access to Justice Technology Bill of Rights (ATJ-TBoR) process (Process), described in detail in both its substantive and procedural aspects in other papers in this volume,¹ has the potential to have a major impact on access to justice in the state—its first and primary goal. In addition, however, it has the potential to have broader implications in the legal world, in the process of legal innovation, in access to other services, and internationally. This paper is intended to start the debate about these implications and how they can be optimized and maximized.

I. THE LESSONS OF THE ATJ-TBOR PROCESS

In the now almost three years of the Process, we have struggled to create a document that would be broadly legitimate, that would be adopted in a manner that would give it real authority and force, and that would be concretely useful in shaping the future and resolving disputes about that future. Some of the techniques utilized in the Process may well have applicability in other endeavors.

A. General Principles

Perhaps the most critical decision made in the development of the Principles of the Access to Justice Technology Bill of Rights (Principles) was to focus early on general principles. The benefit to generality is that it allows agreement on fundamentals even among those who may disagree strongly about detail. Indeed, it turned out during the

---

* Richard Zorza, Esq., A.B., Harvard University, 1971, J.D., Harvard University, 1980, has been the primary consultant to the Access to Justice Technology Bill of Rights process.

drafting and internal comment process that it was much easier to get agreement about general principles than it would have been to get such agreement on specifics.

These general principles can now play at least three roles. First, they can be used to guide the many bureaucratic and institutional processes inside and outside the court system that must wrestle and are already wrestling with issues relating to the use of technology. Second, they are guiding the more specific processes such as the development of Promising Practices. Finally, when there is a disagreement about specifics, the general principles can be the framework for an intellectually coherent analysis of these specifics. The true test of the general principles will be whether they have enough intellectual power to guide these processes and resolve disputes in a way that leads to results that are legitimate for all sides.

During the extended public but internal comment process, there was much criticism of this generality. Commentators feared that the lack of specificity would render the document useless as a guide to action. As a general matter, these comments tended to be associated with particularly strong views about particular outcomes—especially in the area of privacy. There was fear, for example, that the absence of specificity with respect to how privacy and openness issues should be balanced, and the absence of listing of particular information that should be protected, would make it impossible to use the Principles to protect privacy. The test of the general approach will be whether in the process of long-term implementation the general language is given specific content, or whether the good intentions of the document are watered down to nothing. This much is sure, however: the document would never have stood a chance for acceptance had it taken specific stances on currently controversial issues. An excess of specificity in the privacy area might well, for example, have triggered fears of excessive restriction of information. The generality of the Principles is also required by the need for them to remain appropriate even as technology itself changes, and change it does, with extreme speed.

2. The Promising Practices Committee and Promising Practices tool are also known as the Best Practices Committee and Best Practices tool in Washington State. See infra note 14.

3. Reference here is to the many stages of comment that the document went through within the Process, not any comment process for formal legal adoption. See Horowitz, supra note 1.
Reflections on the ATJ-TBoR Process

B. Mooting and Testing

Given the risks of overgenerality, it was important to test the viability for the general principles in real world situations. Those engaged in the drafting process for the Principles found one of the most useful tests to be a “mooting” of a hypothetical against the then-draft Principles of the ATJ-TBoR. In early 2002, we designed an electronic filing hypothetical, envisioned a litigation brought against this planned electronic filing project, and recruited advocates to argue for and against the plan. There was no lack of specific argument on the facts of the hypothetical, and the general language of the then-draft of the Principles. In particular, the oral arguments focused the drafters’ attention on the mandatory or hortatory nature of certain language and the relationship to local court rules, both of which insights proved highly useful to the drafting team. More generally, however, given how helpful the “mooting” was, the drafters were left with a sense of astonishment that such a testing process is not routine in drafting legislation or rules.

4. An edited text of the hypothetical (which has also been used in CLE programs) follows: “A county is about to launch an electronic filing system. The system would: include electronic filing over the internet; set up computers in the courthouse so that people can use them to initially file and subsequently add to cases if they do not have access to the network; set up an agreement with the copy shop so that people can use it; require an e-filing fee, which MUST be paid at filing for all documents (this is small—$5.00); specify that if you use the copy shop, there is a $15.00 additional fee; allow the public to use the Internet to search through all documents that are electronically filed; charge a fee for this searching, 50% of the fee will go to the court and 50% to the private vendor; still allow paper filing; include the possibility that paper filed cases that do not require an early in-court hearing may have a lower handling priority relative to cases on the electronic docket; provide that paper files CANNOT be searched, except manually as in the current system; offer software that includes chat and audio and videoconference features; and includes no standards as to when these features might be used. In addition: the software that is planned for filing is in use in one other state. In that state, the local consumer group claims that it is hard to use for people with less than a tenth grade education; the software is hard for electronic ‘readers’ to use and has been criticized by advocates for the blind and visually impaired in other states; the software does not include document assembly features—that is to say the litigants have to generate their own documents, either using a word processor, or by purchasing or obtaining a form; the court provides some forms online, and ones in paper.”

5. The specific “mooting” assumed an action being brought against a court to enjoin the electronic filing system. The use of that procedural content at that point in the drafting process should not necessarily be construed to imply that the current draft envisions such a procedural route. See infra Part I.F (discussing enforcement procedures).

C. Maximizing Participation and Legitimacy Through Many Related Processes

As is clear from other papers in this publication that describe at least in part the different processes in which the project engaged, each process provided valuable insights and ideas. What may be less obvious from these narratives is the way that each of these processes built on each other. The surveys conducted by the ATJ-TBoR Judiciary and Court Administration Committee helped to alert judges and administrators to the issues that were faced by their colleagues and that informed the drafting of the Principles document itself. The focus groups conducted by the ATJ-TBoR Outreach Committee set an important context for the Principles and for the issues faced by the ATJ-TBoR Implementation Committee. The design process engaged by the ATJ-TBoR Opportunities, Barriers, and Technology Committee highlighted the possibilities and the risks, further illuminating the drafting process. Finally, and perhaps most significantly, the ATJ-TBoR Promising Practices Committee demonstrated the value of the intellectual structure established by the draft of the Principles as it used that structure to establish promising practices that would give concrete meaning to the idea of the Principles.

This system of interlocking committees meant that all participants contributed to the ultimate product and came to feel a sense of ownership. That is an important lesson for future such projects, particularly those that are as ambitious in their reach.

---


8. These surveys, conducted relatively early in the process, made clear the penetration of technology into day-to-day judicial court processes, the level of interest in the issues raised by technology and access to justice, and highlighted the need for a systematic approach. The results of the survey will be made public and posted on the ATJ-TBoR website, http://www.atjtechbillofrights.org, in early March, 2004.


10. T.W. Small et al., supra note 7.
Reflections on the ATJ-TBoR Process

D. Focus Group Work

It is hard to overstate the value of focus groups in a process designed to be so open. The Process used two different sets of focus groups, the first early in the process to identify the needs and perceptions of a variety of groups, mainly the excluded, and the second near the end of the process to identify the impact and potential for those institutions most likely to be affected by the document.

Both processes had the effect of destroying stereotypes and keeping the process focused on the concrete needs of the players. The early focus groups told us something that few knew—that the problems of technology and access to justice are much less problems of access to technology than of the underlying content. The specific summary findings of these groups were as follows:

- With some exceptions, the degree to which technology has penetrated these groups is significant. Most have some basic computer skills, at least occasional access to a computer and the Internet, and some level of comfort with using the Internet (although perhaps 30–35% appear to be left behind).
- In spite of the foregoing, focus group participants are having a hard time finding useful law-related information on the Internet. Among the problems reported were: (a) the available information was too general; (b) there was too much information to absorb; (c) information was promotional, rather than substantive; (d) people have to spend limited available time searching rather than finding useful information; and (e) the language used in law-related material is too technical or not otherwise understandable.
- Except for really important interchanges, participants like the idea of dealing with the justice system remotely by

11. “The six low-income or otherwise vulnerable groups were chosen because it was thought they might have quite different experiences, skills and difficulties with respect to the use of technology. However, although there were differences between the groups [inmates, for example, had more difficulty obtaining access to the hardware], there were generally more similarities than expected.” D. Michael Dale, Outreach Working Committee, Access to Technology Bill of Rights Committee, Washington State Access to Justice Board, Technology and Access to the Justice System: Conversations in Focus Groups with Users of the Justice System in Washington State, in Report of the Outreach Working Committee of the Access to Justice Technology Bill of Rights Committee of the Washington State Access to Justice Board 3–4 (June 30, 2003), available at http://www.atjtechbillofrights.org/tbor/tbordocs/focusgroupsreport.
technological means. Since they are fearful of the courts, they see as a good thing the degree of distance from the courts made possible by technological means of communication. Exceptions to this are that they uniformly dislike voice mail, as well as any other obstacles that prevent them from getting good and timely information.

- Confidentiality and privacy concerns didn’t generally seem to be significant to the low-income groups. The notable exception was the domestic violence group, all of which were women, all of whom were very concerned about privacy and security. To a lesser extent, but an exception nonetheless, was the veterans homeless group. It should be noted that the issue of public on-line availability of case records was not prompted or raised.

- Incarcerated men and women have very little access to any technologies, although they might benefit in appropriate ways if correctional system security concerns could be worked out. The glaring and outstanding exception to those having any measurable degree of access to or familiarity with modern technology is the extent to which farm workers have not been exposed to new technologies, do not understand them, and cannot yet use them. If they are to be integrated into a justice system that relies to any significant degree on technology, farm workers and others similarly situated (such as persons with language or cultural barriers, or who live in remote locations) will, besides needing access to the technology itself, require either intensive training and assistance, parallel provisions for access that rely on more traditional, non-technical means of access to information, or perhaps both. Incarcerated persons, both male and female, irrespective of the kind of institution, are less of an exception, but are an exception nevertheless. The women seemed generally more technology savvy than the men, who were at best mixed in their knowledge or skills.\(^\text{12}\)

The second set of focus groups, conducted by the Implementation Committee, told us that most institutions are much less nervous about being asked to comply with underlying values than many of their
Reflections on the ATJ-TBoR Process

bureaucratic leaders assume.\textsuperscript{13} The Process would not have achieved the legitimacy it did achieve without these focus group processes.

\textbf{E. Promising Practices}

The Promising Practices process was designed to generate a product that would complement the general Principles and would fill in those principles with concrete, specific, and situation-tailored recommendations. These Promising Practices would have substantially less legal force than the Principles themselves, and their concreteness would not raise the political problems that too much specificity in the underlying Principles would have raised.

The process of drafting these Promising Practices—which is now funded in part by the State Justice Institute (SJI)\textsuperscript{14}—still continues, but the documents already in draft show that they can be both specific and broadly useful. The integration of these Practices, in cooperation with NP\textsuperscript{15}ower, the Seattle based national network of nonprofit technology support organizations,\textsuperscript{15} into its TechAtlas tool,\textsuperscript{16} and the inclusion of resources and tools with the recommendations, make them a far more effective tool for decision makers.

More generally, however, the Promising Practices process and product validate the power of Principles approach. The Promising

\textsuperscript{13} The ATJ-TBoR Implementation Committee focus group report results will be made public and posted on the ATJ-TBoR website, http://www.atjtechbillofrights.org, in early March, 2004.

\textsuperscript{14} SJI Grant No. SJI-03-N-108, \textit{ATJ-TBoR Best Practices Template}, at http://www.statejustice.org/grantinfo/apptech.htm (last visited Nov. 24, 2003). The version intended for national use, as funded by SJI, is termed “Promising Practices,” to indicate their lesser force for states other than Washington. This is appropriate, given the far higher involvement of multiple stakeholders from Washington State, and the extent to which the recommendations come from experience and lessons learned in this state.

\textsuperscript{15} NP\textsuperscript{15}ower describes itself as follows on its webpage:

NP\textsuperscript{15}ower helps nonprofits use technology to expand the reach and impact of their work. NP\textsuperscript{15}ower is a federation of independent, locally based nonprofits providing accessible technology help that strengthens the work of other nonprofits. . . . [T]he NP\textsuperscript{15}ower Network now has roots in ten communities and reaches thousands of nonprofits in over 50 communities nationwide.

\textit{NP\textsuperscript{15}ower Home}, at http://www.NP\textsuperscript{15}ower.org (last visited Nov. 24, 2003).

\textsuperscript{16} The TechAtlas tool is described as “a web-based planning tool that your nonprofit can use to assess your current technology use and to receive recommendations on how to better implement technology to achieve your mission.” \textit{TechAtlas}, at http://www.techatlas.org/tools/origins.asp (last visited Jan. 13, 2004). The tool in its general form is available at http://www.techatlas.org/tools (last visited Jan. 13, 2004). As adapted for use by the Process, TechAtlas will ask users a number of questions about any chosen technology implementation and whether it meets the Promising Practices for that implementation, and then provide concrete suggestions for how to meet any that it does not yet meet.
Practices were generated by applying those general Principles to each of the substantive access technology areas for which Promising Practices are to be developed. In each case, the committees asked what concrete attributes the technology needed to have to meet the goals of the Principles.

That this process seems to have worked well confirms the validity and power of the Principles. If they work in this context and in this time they should work in other contexts and other times. A similar process of using the Principles as a guide to what should be done has been followed in developing Giving Life to the Access to Justice Technology Bill of Rights, a document designed to foster involvement in implementation of the vision by as many institutions as possible.17

F. Issues of Mandate, Guidelines, and Enforcement Mechanisms

By far the most difficult question in the entire process has been the question of how to structure and position the ATJ-TBoR so that it has the appropriate legal force. The challenge is that without clear legal authority, there is a great danger of the critics being right—of the document merely gathering dust on a shelf. On the other hand, the greater the authority and enforceability of the document, the greater the risks that stakeholders will fear an intrusion into their traditional areas of authority and discretion and that particular groups will see the enactment as inconsistent with a previously taken highly specific position.18

In coming up with a proposed resolution that the drafters believe satisfies all these concerns, they were guided by a number of core insights: (1) that the ATJ-TBoR must have sufficient authority to be given real effect in the real word; (2) that it deals with access to justice, and therefore the source of that authority should be the State Supreme Court in its rulemaking authority; (3) that the enactment cannot practically be seen as changing existing substantive law or creating on its own new causes of action; and (4) that the impact of the Principles will play out in a number of different forums and ways, including how the complex administrative structures of the justice system deal with the design, deployment, and use of technology, how the rulemaking

17. To view a draft version of the Principles, see Washington State Access to Justice Technology Principles, supra note 1.

18. The risk of this last danger is probably greatest in the area of privacy and openness, in which there is already extensive law and extensive dispute.
Reflections on the ATJ-TBoR Process

processes of the justice system deal with issues of technology and access to justice, and how the day-to-day business of the courts is conducted.

The solution, which may be termed “mandated consideration,” requires the bodies under the authority of the Washington State Supreme Court, including courts conducting their day-to-day business, to consider, together with other rules and governing law, the ATJ-TBoR. Given the general nature of the document, mandated consideration is unlikely to dictate the details of the result of that consideration. But it does mean that a refusal to consider the ATJ-TBoR, or a blanket ignoring of its Principles, would violate the proposed court rule. In other words, the matter is one of sound discretion.

Similarly, the Principles do not create, on their own, new causes of action. Thus, a claim could not be predicated solely on an alleged violation of the specifics of the Principles. But the Principles are very much proposed to be a court rule that can be used in interpreting the relevance, scope, and requirements of any other law relied upon.

In the event that the mandated consideration approach is insufficient, the various ongoing implementation processes of the ATJ-TBoR will bring to decision makers such as the court or the legislature options for making sure that the vision in the document is complied with more fully. 19

II. IMPLICATIONS FOR THE UNITED STATES LEGAL WORLD

The resonance that the Principles have achieved in Washington State suggests that they are relevant to the broader United States legal world. While Washington State’s status as a tech “early adopter” may have resulted in the issues we have addressed coming more quickly to the fore, neither their relevance nor their potential force throughout the United States can hardly be doubted.

A. The Potential for Broader Adoption

There have been expressions of interest in adopting the Principles in several states. The speed of any such adoption will depend on the

19. At least as a theoretical matter, approaches might include setting up a separate reporting and monitoring system, strengthening the language of the rule to require more than “consideration,” establishing a cause of action, or even placing enforcement jurisdiction in a specific court. No opinion is here implied as to efficacy or appropriateness of such approaches.
success of the enactment in Washington State, the extent of its impact, and the alliances it succeeds in creating.

At the federal level, the current political gridlock would appear to make progress impossible except possibly in a narrow and limited sense within the federal courts. A possible, and perhaps narrower, federal court rule enactment (subject to Congressional review) would at least guide federal court automation activity.

B. Impacts on More General Adjudication of Technology Issues

In the process of the Principles, there has as yet been little attention given to the potential impact on judicial sensibilities. The Principles have had a major educational impact on judges, increasing their understanding of the choices made about technology, the impact of these choices, the interrelationship of these choices with the accessibility of institutions, and any cost issues. This impact has been enhanced, of course, by direct judicial participation in the many processes of the Principles.20

It is perhaps inevitable that judges will bring this heightened sensitivity to other cases that have nothing to do with the courts. One possible example: a few years ago, there was extensive litigation about the detail required in welfare termination notices. The staple government defense—the “Fortran defense”—was that it was too hard to reprogram the computers. Judges may well be more skeptical of such claims now, given their own experience with keeping technology under the control of guiding values.

C. Constitutionalizing Access to Justice

More ambitious, however, is the possibility that the Principles can play a role in a broader re-conceptualization and constitutionalization of access to justice. On the litigation front, the attempt to create a “civil Gideon”21 has been largely dead-ended at the federal level for over

20. There has been judicial participation in almost every committee and workgroup including the Drafting Group, the Promising Practices Committee, and the Judiciary and Court Administration Committee.

21. The relationship of the concept of a “civil Gideon,” expanding the right to counsel to civil cases, to other areas of innovation is discussed in Richard Zorza, Discrete Task Representation, Ethics and the Big Picture, 40 FAM. CT. REV. 19, 19–20 (2002).
Reflections on the ATJ-TBoR Process

twenty years and at the state level for over a decade, although a few states, including Maryland and Washington, are currently generating new efforts. Similarly, at the federal constitutional level the attempt to pull down financial barriers to fees for access has come to a halt, although on the state-by-state level, there are broad systems of waiver in place. All the debates, however, have taken place with an assumption of stasis. Access has been assumed to be constant in technique and methodology and constant in cost.

How technology can change this picture—and therefore the balancing test in even existing case law—is through a radical reduction in the costs of providing access and a dramatic increase in the methods of providing access. Buried in most due process analysis, sometimes explicitly, is the financial burden on the state of providing process.

As we have found in the Process, appropriately deployed technology can create low-cost and effective routes to access that are far cheaper than the traditional full representation of a lawyer. This makes it possible for us to argue for a right of access. In other words, people who need access to justice are entitled not to a lawyer, but to whatever it is that they need to obtain proper access.

Technology radically changes that calculation and makes such a broad claim more acceptable. Moreover, as the Principles gain broader

---

22. In Lassiter v. Department of Social Services, the U.S. Supreme Court rejected the blanket claim of federal constitutional right to counsel in parental termination proceedings, and stated:

The case of Mathews v. Eldridge . . . propounds three elements to be evaluated in deciding what due process requires, viz., the private interests at stake, the government’s interest, and the risk that the procedures used will lead to erroneous decisions. We must balance these elements against each other, and then set their net weight in the scales against the presumption that there is a right to appointed counsel only where the indigent, if he is unsuccessful, may lose his personal freedom. 452 U.S. 18, 27 (1981) (citation omitted).


24. See O’Connor v. Matzdorff, 76 Wash. 2d 589, 600, 458 P.2d 154, 160 (1969) (“[C]ourts have found within their powers an inherent power to waive the prepayment of court fees, where a suitor or defendant has shown that he is impoverished, regardless of statutory authority.”); see also Hous. Auth. v. Saylor, 87 Wash. 2d 732, 742, 557 P.2d 321, 327 (1977) (“[C]ourts retain an inherent power to waive their own fees in order to consider a case where it is made to appear that justice requires it.”); Neal v. Wallace, 15 Wash. App. 506, 508, 550 P.2d 539, 541 (1976) (“Waiver of fees for an indigent is a discretionary act within the inherent power of the court.”). Several of the above cases list other Washington cases considering indigence in the context of waiving fees.


26. Examples are laid out in Horowitz, supra note 1.
effects in the justice system, and as more access technologies are developed and deployed, the power behind this argument will become increasingly greater and the potential for establishing a broader right of access will become more realistic.27

III. IMPLICATIONS BEYOND THE LEGAL WORLD

The organizers of the ATJ-TBoR process have always believed that the process would have applications and implications in other areas, particularly in other service areas, such as health, senior, and consumer services.

A. The Relationship of Technology Innovation and Access to Services

Technology Creates New Routes to Access

The obvious lesson from our work is that technology creates, and can be shaped to create, new broadly defined routes of access to services. This means that in any area—medical, communications, financial planning, etc.—technology creates new ways of delivering the services, and therefore new routes to access the services.

Technology Changes the Politics of Any Market

Technology challenges all the political and market configurations and compromises that lead to stasis, and creates an environment in which fundamental questions can be asked about access, services, and equity, and how the levels of access, services, and equity can be changed. These questions can then be resolved in whatever the appropriate forum for that area is—the market, the regulatory environment, or some combination of the two.

Technology Offers Transparency

One of the most powerful of the lessons from the ATJ-TBoR is the power of transparency—the idea that with technology it is much easier to know what is going on. Data is easily susceptible to analysis and openness. This changes the equation with respect to power over information and thus over the ability to manage change.

Technology Changes the Roles of Capital, Investment, and Standardization

Technology in any area requires capital. This means that there is a pressure to standardize, to enhance the power of larger players, and to plan and build for the long term. This creates larger

Reflections on the ATJ-TBoR Process

markets, lesser short-term and greater long-term risk of monopoly. Also, this increases the power of a small number of large players in determining access to services.

B. The Value of a Similar Process in Such an Environment

All these forces increase the power of, and the need for, an ATJ-TBoR Process in other realms. Without such a process, the risk of surrender to the force of change is greater. With such a process, the best thinking can control that process of change.

Technology Creates the Possibility of Alliances Since All Can Identify with Risks. One of the unique aspects of technological change is that anyone can fear it. There is no one so well-connected or so wealthy that they cannot imagine being excluded by inability to use a technology. The result is that it is possible to build a much broader access coalition around issues of technology than around poverty, class, or even age or other suspect classifications. This is as true in the impact of technology in access to health or housing services as it is to justice.

The Value of a Principles-Driven Approach to Optimizing Opportunities. Similarly, the power of a principles-driven approach is particularly appealing in responding to technology-driven changes in any area, not just justice. Principles allow individuals to feel that they have regained power over something that they do not understand, something that might be out of control, something unpredictable. This is what they feel about technology. The principles themselves may vary with the substantive area, although the general issues addressed will all have power in any area and provide a powerful starting place.

The Value of an Access Approach. Of similar use is the general appeal to access as a value. It is not a technology justice bill of rights—a document that would have thrown us into the middle of a number of irresolvable disputes about technology. Rather, it is about access, a formulation that is less controversial, safer, and more universal. Buried, of course, in the word “access” are a number of possible meanings. Strictly speaking, access to justice can mean access to justice services, access to just results, or even access to a just society.

The Force of Transformative Opportunities. The key to creating consensus and energy around such a process is spreading the idea that technology is not about re-slicing the cake, but about finding new ways to transform the size and potential of the cake. Then all stakeholders can see that they can gain.
IV. IMPLICATIONS FOR AN EMERGING INTERNATIONAL RIGHT TO ACCESS TO JUSTICE

It is seldom recognized in the United States that most industrialized countries have a far stronger commitment to access to justice than does the United States, at least if that commitment is measured by a country’s willingness to pay for legal help for those who cannot afford it. The United States pays far less than other countries for legal aid services, even on a per capita basis. The figures are even more dramatic when corrected for gross national product (GNP), legal GNP, or poverty population, or the number of problems that are dealt with in the respective legal system.28

Even more interestingly, the current Draft European Union Constitution includes under the heading “Right to an effective remedy and a fair trial” the following language: “Legal aid shall be made available to those who lack sufficient resources insofar as such aid is necessary to ensure effective access to justice.”29 This is at least in part a codification of a 1979 decision, Airey v. Ireland.30 The Airey decision dealt with the right to counsel and was decided under Article 6, Section 1 of the European Convention for the Protection of Human Rights and Fundamental Freedoms,31 which states: “In the determination of his civil rights and obligations . . . everyone is entitled to a fair and public hearing . . . by an independent and impartial tribunal.”32 Similarly, the Canadian Supreme Court held in 1999 that there was a right to counsel when the state attempted to take even short-term custody of a parent’s child.33


31. Id. at 313.


33. See New Brunswick v. J.G., [1999] 3 S.C.R. 46. J.G. relied upon Section 7 of the Canadian Charter of Rights and Freedoms, which states that “[e]veryone has the right to life, liberty and security of the person and the right not to be deprived thereof except in accordance with the principles of fundamental justice.” Id.
Reflections on the ATJ-TBoR Process

In the United States, as in Europe, the potential of technology and of an access to justice technology lies in the words “necessary to ensure effective access to justice.” The perspective of the Principles raises the possibility of “effective access to justice” by making clear how realistic such access is. At the same time, it lowers the cost and increases the feasibility by reducing the range of circumstances in which there is a need for an attorney to “ensure effective access.”

In the European context, this insight may make it less financially frightening for the language to be adopted, and when adopted, to be followed and broadly construed. Within the rest of the world, the cost-lowering and access-creating possibilities of technology should make similar enactments more appealing and acceptable. The Principles might well provide an opportunity and organizing tool for groups in various countries committed to increasing access to justice. They might find allies, including in the technology industries, parallel to those recruited by the Washington State process.

All these rights derive ultimately from Article 8 of the United Nations Universal Declaration of Human Rights, which mandates that “[e]veryone has the right to an effective remedy by the competent national tribunals for acts violating the fundamental rights granted him by the constitution or by law.” This common heritage and language should encourage advocates to think about how to create an international right of access to justice, and about the relationship of that right to the appropriate use of technology.

V. HOW WASHINGTON STATE CAN MANAGE ITS IMPLEMENTATION OF ATJ-TB OR TO FACILITATE BROADER USE OF THE MODEL

While our first obligation is to make sure that the Principles work as well as possible for the people of Washington State, our opportunity is also to move forward in this state in ways that increase the chances of the document having the broader impact outlined above. Below are some thoughts that the enactors and implementers may well bear in mind as the process moves forward.

34. For a review of non-U.S. attempts to create non-attorney methods of access, see generally Tiffany Buxton, Note, Foreign Solutions to the U.S. Pro Se Phenomenon, 34 CASE W. RES. J. INT’L L. 103 (2002).
A. The Importance of the Adoption Process

A smooth and unified adoption process is crucial for the credibility of the process and the product. That the ATJ-TBoR is seen at the same time as making a difference, yet arousing no focused opposition, will make it far more acceptable in other forums and fields.

B. Issues of Costs and Process

Equally important is management of the costs issue. The main fear that the ATJ-TBoR evokes is that institutions will be burdened with unacceptable “unfunded mandates.” To the extent that the wording, adoption, implementation, and Promising Practice processes can show that the costs issue is manageable, and that the ATJ-TBoR provides a focus for generating agreement on appropriate investments without triggering unacceptable conflict, a powerful stepping stone will have been laid for broader deployment of the core ideas and approach.

C. Showing the Value of Principles-Driven Approach

There has been criticism of the broad principles approach. To the extent that the implementation process can show the value of this approach, that it can show that the principles provide a way of moving forward and structuring consensus even on difficult issues, the power of the whole concept will be validated. This can best be done by using the Principles in every stage of implementation.

D. Importance of Enforceability

Probably the most difficult drafting issue for the ATJ-TBoR has been the enforceability issue. The credibility of the project requires finding the right formula that puts real governmental authority behind the ATJ-TBoR, without undermining the current consensus behind its ideas.

E. Alliances and Engaging the Private Sector

Finally, the continued alliance building process is crucial to the long term power of the concept. Not only will the alliances built make the implementation more powerful, but each state ally can engage its own national and international network in support of broader deployment of the ATJ-TBoR.
Reflections on the ATJ-TBoR Process

VI. CONCLUSION

What has been built in Washington State with the Process is both powerful and unique. With skill and commitment it can have a powerful effect well beyond the state’s borders in increasing access to justice, in transforming our understanding of the relationship between technology and access to justice, and in broadening our understanding of the potential for change to be channeled in the public interest. The judicial and legal leadership of the state, as well as the many community groups with which they have worked on this project, will be able to look back with pride upon a transformative and ongoing initiative, which is likely to have influence far beyond its borders.
Abstract: Federal courts in the United States have embraced electronic access to court records because it promises to allow courts to run more efficiently. At the same time, critics worry that electronically available court records might provide identity thieves with a trove of clients’ personal information. The United States Judicial Conference has adopted a policy endorsing electronic access to court records, but the policy does not contain an express enforcement mechanism to protect clients’ privacy. While court-directed protections, such as Rule 11 sanctions, might help prevent identity theft, they will not help clients after the crime occurs. To recover their losses from identity theft, clients might seek recovery from the attorneys who caused their losses by failing to redact their personal information from court filings. This Comment proposes that clients use the Judicial Conference’s policy as evidence of their attorneys’ duty to redact and demonstrates how clients can prove that their attorneys’ failure to redact was the proximate cause of their losses resulting from identity theft.

Fueled by the rapidly increasing availability of private information, identity theft is one of the fastest growing crimes in America. In an ironic twist, by making court filings available electronically, the federal court system may be unwittingly facilitating identity theft. In 2001, the


Judicial Conference Committee on Court Administration and Case Management on Privacy and Public Access to Electronic Case Files (Judicial Conference Committee) officially recommended that civil records in United States district courts be made available electronically, with limited exceptions, “to the same extent they are available at the courthouse.”

The Judicial Conference Committee hailed electronic access to court documents as a means to improve the efficiency of court operations. However, critics of electronic access have argued that imaged case files will place clients’ personal information in the hands of identity thieves. Responding to these concerns, the Judicial Conference Committee adopted a national policy (Policy) establishing basic guidelines for district courts to follow in determining what documents can be posted electronically and how district courts should protect private information contained within them. The Judicial Conference Committee supplemented the Policy with a Frequently Asked Questions (FAQ) publication designed to answer common questions from court clerks and the bar. Both federal district courts in Washington State have adopted...

3. The Judicial Conference Committee’s recommendation also discussed bankruptcy, social security, and criminal cases. The main topic of the recommendation involved civil cases, however, and this Comment focuses only upon civil cases. See U.S JUDICIAL CONFERENCE, REPORT OF THE JUDICIAL CONFERENCE COMMITTEE ON COURT ADMINISTRATION AND CASE MANAGEMENT ON PRIVACY AND PUBLIC ACCESS TO ELECTRONIC CASE FILES, at http://www.privacy.uscourts.gov/Policy.htm (Sept. 2001) [hereinafter POLICY].

4. See id.

5. See id.


7. POLICY, supra note 3.

Electronic Court Records & Malpractice

the Policy’s recommendations on how to protect clients’ privacy without expanding or elaborating on the recommendations.9

One of the Policy’s key provisions requires attorneys to redact their clients’ private information from all court filings—paper or electronic.10 Yet, the Policy contains no mechanism clients can use to enforce this requirement.11 Therefore, clients may have to rely on Washington’s legal malpractice cause of action to make this new duty meaningful. By suing their attorneys for malpractice, clients can shift their losses back to their attorneys, who were responsible for releasing the clients’ private information.

This Comment explores the sudden rise of electronic access to federal court records and argues that clients can successfully sue their attorneys for malpractice when they lose their identities due to their attorneys’ failure to redact the clients’ personal information from court filings. Part I outlines the scope of the public’s right to access court records, the federal judiciary’s decision to allow electronic access to those records, and the growing crime of identity theft. Part II explains how the Policy attempts to manage electronic access to court records. Part III examines two elements of a malpractice claim under Washington law: breach of duty and proximate causation. Part IV argues that Washington courts should allow client-plaintiffs to introduce the Policy as evidence of their attorneys’ specific duty to redact sensitive personal information from federal court filings. Part IV also argues that attorneys’ failure to redact is the proximate cause of their clients’ losses resulting from stolen identities in situations where the thief obtained the client’s information from the unredacted court filings. Finally, Part IV argues that the criminal act of identity theft is not a superseding cause of clients’ losses.


10. See POLICY, supra note 3.

11. See id.
I. ELECTRONIC ACCESS TO CASE FILES ENDANGERS LITIGANTS’ IDENTITIES

Courts in the United States recognize that the public has a right to access court records. This right can be limited, however, when the public requests a form of access that might lead to improper use of the information contained in the court records. The Policy reflects the federal courts’ decision to allow electronic access to court records. Although this access may be more expansive than the law requires, it could greatly improve the efficiency of court operations. While electronic access to court records has its benefits, critics have raised concerns that such access could facilitate theft of clients’ identities.

A. The Public Has a Right To Access Court Documents, But Not Necessarily To Do So Electronically

There is a strong tradition in the United States of allowing expansive public access to court records, including all official case filings and exhibits admitted at trial. Courts have rejected even arguably admirable attempts to protect the privacy of victims of violent crimes like rape. Nevertheless, the U.S. Supreme Court has held that courts may limit the

---

13. See infra notes 20–24 and accompanying text.
14. See POLICY, supra note 3.
15. See infra note 30 and accompanying text.
16. See supra note 6 and accompanying text.
17. See Nixon, 435 U.S. at 597; Cox Broad. Corp. v. Cohn, 420 U.S. 469, 492 (1975); Ex parte Drawbaugh, 2 App. D.C. 404, 407 (1894) (noting that “any attempt to maintain secrecy, as to the records of the court, would seem to be inconsistent with the common understanding of what belongs to a public court of record”). But cf. Seattle Times Co. v. Rhinehart, 467 U.S. 20, 33–34 n.19 (1984) (discussing the traditional common law understanding that items not filed during a case were not subject to public disclosure and that trial courts historically exercised control over access to courthouse records). American courts long ago rejected the English practice of requiring the public to demonstrate a legitimate interest before accessing court records. See Drawbaugh, 2 App. D.C. at 406–07. But see In re Caswell, 29 A. 259, 259 (R.I. 1893) (denying access which is sought simply out of mere curiosity or to create a scandal). This reaction was motivated by the court’s concerns that once the right to access court records was subjected to judicial scrutiny, a Pandora’s box of limitations could evolve. See Richmond Newspapers, Inc. v. Virginia, 448 U.S. 555, 569–81 (1980) (discussing corollary right to access in the context of criminal trials).
Electronic Court Records & Malpractice

form of access to court records where “court files might have become a vehicle for improper purposes.”

In *Nixon v. Warner Communications, Inc.*, the Court recognized a crucial distinction between the right to access and the form of access. In *Nixon*, the Court held that the trial court properly limited the public’s access to paper transcripts of tapes containing the voice of President Nixon that had been admitted into evidence as part of the Watergate criminal prosecution. The Court decided that a paper transcript of the tapes, rather than copies of the tapes themselves, would allow the public access to the actual information contained in the tapes, yet minimize the likelihood that the tapes would be put to improper commercial or spiteful uses. Under *Nixon*, then, courts may regulate the form of access to court records when one particular form of access is less likely to result in improper usage of the records than another.

Congress has encouraged federal courts to make case files available electronically, but it has not mandated any specific level or means of electronic access. Furthermore, while Congress has directed the U.S. Supreme Court to adopt rules of civil procedure to address electronic access, it has not imposed a specific time requirement on the Court. Nevertheless, in 2001, the federal court system’s Judicial Conference

19. *Nixon*, 435 U.S. at 598 (listing improper purposes). While identity theft was not listed as an improper purpose, the Court included actions such as use out of spite or for commercial gain, neither of which is inherently criminal. *Id.*


21. *See id.* at 609.

22. *See id.*

23. *See id.* at 603.

24. *See id.* at 603–09.

25. Congress has not required the federal courts to provide electronic access, but has indirectly approved the practice by authorizing the Judicial Conference Committee to establish user fees for the service. *See Chronology of the Federal Judiciary's Electronic Public Access (EPA) Program, available at* http://pacer.psc.uscourts.gov/documents/epachron.pdf 1 (last visited Nov. 5, 2003). Some uncertainty remains as to whether the courts, the legislature, or both will be responsible for managing electronic access. *See In Depth: Privacy and Electronic Access to Court Files, E-FILING REPORT, 1 No. 2 e-Filing Rep. 14 (Jan. 2001).*

26. *See E-Government Act of 2002, Pub. L. No. 107-347, § 205, 116 Stat. 2899, 2913–15.* Congress has only required that the courts post paper filings if the courts scan those documents electronically; if the courts choose not to scan the documents, they are not required to post them. Congress has also directed the U.S. Supreme Court to adopt rules of civil procedure designed to protect privacy concerns. Congress did not dictate what the new rules should look like or when they must be completed. *See id.*

27. *See id.*
Committee decided to encourage district courts to make court documents available electronically.\textsuperscript{28}

The Judicial Conference Committee decided to embrace an Internet database method of access that provides the public with the same scope of electronic access as that available at the courthouse.\textsuperscript{29} It based this decision largely on the theory that courts will function more efficiently if they provide increased accessibility for members of the public and lawyers who are not in close proximity to the courthouse, reducing the workload in clerks’ offices.\textsuperscript{30} To achieve this efficiency, the Committee envisioned a nationwide system providing extensive access.\textsuperscript{31}

Having decided to make court records available electronically, the Judicial Conference Committee placed only nominal limits on who can electronically access those documents.\textsuperscript{32} Court records are available through the federal courts’ official Public Access to Court Electronic Records system (PACER).\textsuperscript{33} PACER is an Internet-based database accessible virtually world-wide.\textsuperscript{34} The system requires users to input a password to log onto PACER, but almost anyone can obtain a password within a few weeks by request.\textsuperscript{35} The password requirement is designed to create an electronic trail, which the PACER administrators can trace if a problem arises to determine who accessed specific information.\textsuperscript{36} Each district court maintains its own PACER database. While practices vary among the districts, fifty-eight districts allow PACER users to view filings that have been electronically scanned.\textsuperscript{37} The PACER system then compiles limited information about each civil case, including party

\begin{flushleft}
\textsuperscript{28} See POLICY, supra note 3.
\textsuperscript{29} See id.
\textsuperscript{30} See id.
\textsuperscript{31} See id.
\textsuperscript{32} See id. The requirement that the public first obtain a password appears to be the only limit upon access under the Policy. See id.
\textsuperscript{33} See id.; see also PACER, at http://www.pacer.psc.uscourts.gov (last visited Nov. 29, 2003).
\textsuperscript{34} While the primary means of access is now Internet based, a direct dial-in version is also available for those without Internet access. See PACER, supra note 33.
\textsuperscript{35} See id. (explaining that there is no cost to register for the PACER password, and the official registration form only requires users to provide a name and address to which the password can be mailed).
\textsuperscript{36} See POLICY, supra note 3. The ability to trace a user is of course dependant on the user having submitted truthful information on their registration form.
\textsuperscript{37} See PACER Web Links, at http://www.pacer.psc.uscourts.gov/cgi-bin/links.pl (last visited Nov. 4, 2003). Districts that do not yet scan documents generally provide some alternative information, often in the form of the docket information. Id.
\end{flushleft}
Electronic Court Records & Malpractice

names and the subject of each case, into a national database called the U.S. Party Case Index, which is accessible with the same password. This compilation enables nationwide searches for specific kinds of cases across all local PACER systems.

B. Court Records Contain All of the Information Necessary To Steal an Identity

Critics quickly vocalized their concerns about litigants’ privacy following the Judicial Conference Committee’s decision to allow electronic access to court records. The most widespread concern was that criminals would be able to use information contained in court records to steal clients’ identities. Identity theft is one of the fastest growing crimes in America, and several commentators have warned that an official case file includes everything needed to steal an identity. According to the Federal Trade Commission (FTC), all that is needed to steal an individual’s identity is that person’s Social Security number, date of birth, and financial account numbers. The U.S. Party Case Index has simplified access to this information by permitting searches by claim type. A thief need only determine which type of claim would most likely require filings containing the information necessary to steal an identity and conduct such a search.

The consequences of identity theft can be extensive and take on a variety of forms. Victims of identity theft often spend substantial time and money attempting to get creditors to institutionalize victims’ losses. The average victim suffers $4,800 to $10,200 in losses from

39. See id.
40. See Alexander & Slone, supra note 6, at 438; Givens, supra note 6 (noting situation where a singles club accessed divorce records to find marketing targets).
41. See EHLERS ET AL., supra note 6, at § 3.13; Givens, supra note 6.
42. See supra note 1 and accompanying text.
43. See EHLERS ET AL., supra note 6; at § 3.13; Givens, supra note 6.
44. See ID THEFT, supra note 1, at 1.
45. See Nature of Suit, at http://www.pacer.psc.uscourts.gov/natsuit.html (last visited Nov. 5, 2003); U.S. Party Case Index Overview, supra note 38. For example, by entering code 371 into the system to search for Truth in Lending cases, thieves might expect to find more relevant personal and financial data than if they entered code 720 for Labor/Management Relations or searched broadly under all case types.
46. See Christopher P. Couch, Comment, Forcing the Choice Between Commerce and Consumers: Application of the FCRA to Identity Theft, 53 ALA. L. REV. 583, 586 (2002) (citing
identity theft, and spends thirty to sixty hours resolving the matter. Additionally, many victims of identity theft experience substantial indirect costs. These indirect costs include medical expenses for treating theft-induced stress, sleeplessness, and severe depression.

II. FEDERAL COURT POLICY REGULATES THE POSTING AND FORMATTING OF ELECTRONIC DOCUMENTS

When it decided to make electronic court records available electronically, the Judicial Conference Committee proposed a model Policy to be adopted by individual district courts. Many federal district courts, including both Washington State districts, have adopted this Policy. The Policy acknowledges concerns about the risks to litigants’ privacy and attempts to provide a uniform national framework for balancing those risks with the benefits of increased access to court records. Any privacy concerns, however, are tempered by the fact that courts using the PACER system have reported few problems with it thus


47. See FTC—IDENTITY THEFT SURVEY REPORT, supra note 1, at 6. Because of various methods to shift the loss back to creditors, individuals on average pay between $500 and $1,200 of the loss from their personal resources. The out-of-pocket cost differs depending on whether the crime is the misuse of an existing account or the opening of a new account. Id. The legal rights of creditors and a client’s possible contributory responsibility for failing to minimize losses are beyond the scope of this Comment.

48. See IDENTITY FRAUD: INFORMATION ON PREVALENCE, COST, AND INTERNET IMPACT IS LIMITED, supra note 1, at 49; Couch, supra note 46, at 585–86.

49. See Couch, supra note 46, at 586.

50. See POLICY, supra note 3.


52. See POLICY, supra note 3.
far.\textsuperscript{53} Despite this good news, the Policy addresses the continuing risks posed by electronic access by (1) limiting the kind of documents that are available electronically, (2) proposing that trial courts seal certain documents from electronic access, (3) suggesting that trial courts impose Rule 11 sanctions to punish attorneys who use electronic access as a means to maliciously reveal opposing litigants’ private information, and (4) requiring attorneys to redact four kinds of information about their clients.\textsuperscript{54}

\textbf{A. The Policy’s Definition of Case File Limits the Kind of Documents Accessible Electronically}

The Policy’s first means of preventing identity theft is its definition of “case file,” which determines what kinds of documents will be made available electronically. Only items considered part of the case file are subject to imaging and uploading onto PACER.\textsuperscript{55} The Policy defines case file broadly to encompass most litigation-related documents that have traditionally been available at the courthouse.\textsuperscript{56} This definition includes documents submitted by the parties, those created by the court, and transcripts of oral proceedings.\textsuperscript{57} However, the Policy’s definition does not include information such as non-filed discovery materials, trial exhibits not admitted into evidence, and drafts or notes produced by the court.\textsuperscript{58} This definition of case file protects private information by keeping most sensitive information contained in voluminous discovery materials out of public view.\textsuperscript{59}

\begin{itemize}
\item \textsuperscript{53} See id. This conclusion was reached based on an evaluation prior to the advent of widespread imaging by district courts, and its prospective value may be questionable.
\item \textsuperscript{54} See id.
\item \textsuperscript{55} See id.
\item \textsuperscript{56} See id.; Seattle Times Co. v. Rhinehart, 467 U.S. 20, 33 (1984) (noting that “pretrial depositions and interrogatories are not public components of a civil trial,” and that “[s]uch proceedings were not open to the public at common law, and, in general, they are conducted in private as a matter of modern practice”) (citations omitted).
\item \textsuperscript{57} See POLICY, supra note 3.
\item \textsuperscript{58} See id.
\item \textsuperscript{59} See Seattle Times Co., 467 U.S. at 33.
\end{itemize}
Washington Law Review

Vol. 79:407, 2004

B. The Policy Proposes That Judges Seal Files on Request

The Policy also proposes that trial judges use their supervisory power over court records to seal certain documents from electronic access. Before the court will order a file sealed, it must make an affirmative finding, usually upon the request of a party, that the file contains information that should not be available to the public. Clerk’s offices are not required to review filings for sensitive information; it is solely the responsibility of attorneys to review documents for compliance with the Policy. Therefore, courts are unlikely to order files sealed if attorneys, unaware of their obligation to protect clients’ personal information, fail to raise the matter with the court.


The Policy also recommends that courts use Rule 11 sanctions to protect clients. As with the protection offered by sealing documents, the potential protection provided by Rule 11 sanctions largely depends upon attorneys paying attention to the documents filed by both sides. Attorneys who actively review the opposition’s filings will be in a position to inform the court when opposing counsel maliciously files documents containing their clients’ personal information. Less diligent attorneys may not notice when opposing counsel files documents that contain their clients’ personal information until it is too late to prevent

60. See POLICY, supra note 3. The Policy is not clear on what authority or standard exists for the recommended sealing of documents, nor is it clear on whether both the electronic and paper versions will be sealed or just the electronic. Id.; see also E-Government Act of 2002, Pub. L. No. 107-347, § 205(c)(3)(iv), 116 Stat. 2899, 2914 (requiring that if the U.S. Supreme Court’s eventual rules for electronic documents permit the redaction of information, parties be allowed to file unredacted copies under seal with the court).
61. See FED. R. CIV. P. 26(c) (describing the procedure for obtaining protective orders). Because there are currently no standards under the Policy, courts will likely have to rely on the traditional approach under Rule 26(c).
62. See FAQ, supra note 8, at 2.
63. See POLICY, supra note 3. The Judicial Conference Committee, while acknowledging privacy concerns, declined to recommend the adoption of a new rule of civil procedure designed to address the topic. See id. The Committee rejected a formal rule of procedure because adopting such a rule would take several years and privacy concerns needed immediate attention. See id.
64. FAQ, supra note 8, at 2 (noting that clerk’s offices will not review documents).
65. Because the court and the clerk’s office are not responsible for reviewing filed documents the attorney has the sole responsibility to monitor what is filed and take appropriate action. See id.
Electronic Court Records & Malpractice

that information from falling into the wrong hands.\(^{66}\) Rule 11 sanctions, however, are a limited remedy because courts traditionally have been reluctant to impose them in the absence of some action by an attorney clearly warranting judicial sanction.\(^{67}\)

D. The Policy Requires That Attorneys Redact Sensitive Information from Their Case Filings

Finally, the Policy requires that “counsel and pro se litigants” redact, omit, or modify sensitive information from the documents they file with the court.\(^{68}\) There are four types of sensitive information expressly covered by the Policy: (1) Social Security numbers, (2) dates of birth, (3) financial account numbers, and (4) names of minor children.\(^{69}\) This protection places the burden of protecting clients squarely on the party making the actual filing, usually an attorney, and does not require courts to take any specific action.\(^{70}\) Under the Policy, attorneys must “inform their clients that case files may be obtained electronically and ensure private information is not included in the case files.”\(^{71}\) In placing this responsibility on attorneys, the Policy recognizes that “[m]embers of the bar must be educated about the [Policy] and the fact that they must protect their clients.”\(^{72}\) Washington’s federal district courts have heeded this call and taken steps to ensure that the bar is aware of its new duty.\(^{73}\)

\(^{66}\) Whether an attorney’s failure to move to seal or seek Rule 11 sanctions related to an adversary’s filing would support a malpractice claim is beyond the scope of this Comment.

\(^{67}\) See Conn v. CSO Borjorquez, 967 F.2d 1418, 1421 (9th Cir. 1992) (“Rule 11 is an extraordinary remedy, one to be exercised with extreme caution.”) (citation omitted); Operating Eng’rs Pension Trust v. A-C Co., 859 F.2d 1336, 1344 (9th Cir. 1988) (stating that Rule 11 is reserved for “rare and exceptional case[s]”). Because courts have the alternative option to correct the problem by sealing the filing, they may be reluctant to punish attorneys in most situations resulting from simple carelessness.

\(^{68}\) See POLICY, supra note 3. Hereafter, this Comment uses the term “redaction” to represent all three methods of protecting information. The redaction provision of the Policy applies to all documents filed with the courts. See FAQ, supra note 8, at 3. Some district courts such as the Eastern District of Washington have used slightly stronger language like “shall refrain from including, or shall redact” (emphasis added) in their general orders. See General Order, supra note 9.

\(^{69}\) See POLICY, supra note 3.

\(^{70}\) See FAQ, supra note 8, at 2.

\(^{71}\) Id.

\(^{72}\) POLICY, supra note 3.

\(^{73}\) See supra note 9 and accompanying text.
In sum, the Policy primarily places the responsibility of protecting litigants on the judges and attorneys who must develop the actual means for adapting existing procedures to the new electronic reality. The only procedural adaptation the Policy undertakes is its extension of the traditional understanding of what constitutes an official case file. The definition of case file continues the federal courts’ tradition of keeping certain types of information out of the public eye, but it does not decrease the amount of information available to the public. While the Policy suggests three means by which courts and attorneys can protect litigants’ identities, it does not itself create a new cause of action for litigants. Thus, the Policy does not expressly provide clients a means of enforcing one of its central protections: the attorney’s duty to redact the client’s personal information before filing documents with the court.

III. THE CURRENT STATE OF LEGAL MALPRACTICE LAW IN WASHINGTON

Clients who suffer harms as a result of their attorneys’ acts, or failures to act, may seek a remedy through common law malpractice actions against their attorneys. In Washington, a legal malpractice claim consists of four elements: (1) an attorney-client relationship giving rise to a duty of care on the part of the attorney to the client; (2) an act or omission by the attorney breaching the duty of care; (3) damage to the client; and (4) proximate causation between the attorney’s breach of duty and the damages incurred. Once the client establishes an attorney-client relationship, Washington courts have clarified that the remaining elements of a legal malpractice claim are identical to a standard negligence claim.

74. See supra notes 56–57 and accompanying text.
75. This provision simply maintains the status quo protections already recognized for paper files. See Seattle Times Co. v. Rhinehart, 467 U.S. 20, 33 (1984); supra note 59 and accompanying text.
76. See POLICY, supra note 3; FAQ, supra note 8, at 2.
77. A plaintiff would also likely include the identity thief and other responsible entities as parties. This Comment focuses on the action against the attorney because the attorney is likely the only defendant capable of actually paying on a potential judgment. See Benner, supra note 6 (discussing the nature of identity theft and indicating that identity thieves are not usually capable of satisfying any potential judgment).
A. Attorneys Breach Their Duty to Clients by Failing To Act in Accordance with the “Reasonable Attorney” Standard of Care

Legal malpractice decisions focus heavily on the standard of care that attorneys owe their clients, and what constitutes a breach of that standard. \(^{80}\) Washington courts consistently begin their decisions by stating that the applicable standard of care is the “reasonable attorney” standard. \(^{81}\) Washington courts then address what evidence plaintiffs may submit to demonstrate that their attorneys owed them a specific duty under the broad reasonable attorney standard. \(^{82}\) Lastly, Washington courts determine whether the harm alleged is within the general field of harms that the specific duty was meant to protect against. \(^{83}\)

1. The Reasonable Attorney Standard of Care

The Washington State Supreme Court set forth the general standard of care for attorneys in \(\text{Walker v. Bangs}\). \(^{84}\) In \(\text{Walker}\), the court held that the standard of care required by a lawyer performing professional services in Washington state is “‘that de[g]ree of care, skill, diligence and knowledge commonly possessed and exercised by a reasonable, careful and prudent lawyer in the practice of law in this jurisdiction.’” \(^{85}\) This standard of care has been repeated with little comment in several Washington legal malpractice cases. \(^{86}\)

---

\(^{80}\) See \(\text{Hizey, 119 Wash. 2d at 260, 830 P.2d at 651; Marc R. Greenough, Note, The Inadmissibility of Professional Ethical Standards in Legal Malpractice Actions After Hizey v. Carpenter, 68 Wash. L. Rev. 395, 398–400 (1993).}\)

\(^{81}\) See infra notes 84–86 and accompanying text.

\(^{82}\) See infra notes 87–99 and accompanying text.

\(^{83}\) Washington courts to date have not addressed the impact of foreseeability of harm as a factor in determining the duty owed by an attorney. The only Washington legal malpractice case to focus on foreseeability of harm was \(\text{Bullard v. Bailey}\), which did so exclusively in the context of legal causation. 91 Wash. App. at 759, 959 P.2d at 1127. Despite this decision, the general consensus is that the better approach is to treat foreseeability as an aspect of whether the defendant owed the plaintiff a duty. See \(\text{Rikstad v. Holmberg, 76 Wash. 2d 265, 268, 456 P.2d 355, 358 (1969); DiWOLF & ALLEN, supra note 78, § 4.26, at 121. This general approach taken in Washington negligence law should apply equally to legal malpractice cases. See supra note 79 and accompanying text.}\)


\(^{85}\) See \(\text{id. (quoting Cook, Flanagan & Berst v. Clausing, 73 Wash. 2d 393, 395, 438 P.2d 865, 867 (1968).}\)

2. **Plaintiffs Must Demonstrate That Their Attorneys Had a Specific 
Duty To Refrain from the Allegedly Negligent Conduct**

In addition to showing that their attorneys failed to live up to the 
general reasonable attorney standard of care, plaintiffs in legal 
malpractice actions must also prove that their attorneys breached a 
specific duty owed to them. A specific duty is a particular action that a 
defendant should perform in a particular situation. A plaintiff proves 
the attorney breached the standard of care by proving that the attorney 
failed to perform a specific duty. Specific duties usually relate to the 
attorney’s trial tactics and procedures.

To prove a violation of a specific duty, plaintiffs in legal malpractice 
claims must first submit evidence demonstrating that the alleged duty 
existed. In *Hizey v. Carpenter*, the Washington State Supreme Court 
held that plaintiffs may not introduce the Rules of Professional Conduct 
(RPCs) as evidence of their attorneys’ duty, and that the RPCs may not 
be referred to in expert testimony or jury instructions. The motivation 
behind the court’s holding was its reluctance to allow the RPCs, which 
were created by judges rather than a legislative body, to establish the

---

87. See *Rikstad*, 76 Wash. 2d at 268, 456 P.2d at 357. While not a legal malpractice case, the 
*Rikstad* decision did note the general rule that plaintiffs in negligence actions must demonstrate “a 
duty upon defendant[s] to refrain from the complained-of conduct.” *Id.*

88. See Karen J. Feyerherm, *Recent Development, Legal Malpractice—Expansion of The 
(discussing the development of an attorney’s “specific duty to refer” a client to a specialist in certain 
situations); see also *Schooley v. Pinch’s Deli Market, Inc.*, 134 Wash. 2d 468, 478, 951 P.2d 749, 
754 (1998) (stating that a harm must not only be within a general duty to be recoverable, but that 
courts must look to whether the harm was “within the general field of danger covered by the 
specific duty owed by the defendant”); *Raider v. Greyhound Lines, Inc.*, 94 Wash. App. 816, 819, 
975 P.2d 518, 519 (1999) (noting that while a business has a general duty to protect its invitees from 
reasonably foreseeable criminal conduct by third parties, “the harm must lie within the general field 
of danger covered by [the] specific duty owed by the defendant”).

89. See *supra* note 88 and accompanying text.

90. See *Walker*, 92 Wash. 2d at 858, 601 P.2d at 1282 (discussing the need for an expert to testify 
on the specific trial techniques and procedures that should have been applied). However, not all 
legal malpractice cases allege a specific duty related to the trial. *See Hizey*, 119 Wash. 2d at 258–60, 
830 P.2d at 650–51 (discussing the scope of an attorney’s duty not to enter into conflict-of-interest 
specific duty on an attorney to prevent his unlicensed associate from representing himself as an 
attorney to the firm’s clients).

91. See *Hizey*, 119 Wash. 2d at 259, 830 P.2d at 652 (attempting, but failing, to get the RPCs 
admitted as evidence of the attorney’s specific duty to the plaintiff).


93. See *id.* at 265, 830 P.2d at 654.
standard of care for an attorney. The Hizey court provided three reasons why the RPCs should not be formally admitted as evidence of the standard of care: (1) the RPCs expressly deny creating a standard for civil liability for lawyers; (2) allowing the RPCs to serve as evidence of specific duties owed to clients would cause attorneys to minimize the equally important role the RPCs play in protecting the court system; and (3) the RPCs cannot illuminate the specific duties of a reasonable attorney because they outline a vague ethical scheme rather than particular duties owed in particular situations. Washington’s approach of barring the RPCs from legal malpractice actions is one of the most extreme responses, although a majority of states impose some limit on the use of ethical conduct rules in legal malpractice cases.

3. Attorneys’ Duty Is Limited To Preventing Foreseeable Harms

In addition to requiring proof of the specific duty owed, Washington’s negligence law further limits the scope of a defendant’s duty to the prevention of only those harms that were foreseeable. The
Washington State Supreme Court established this requirement in Rikstad v. Hokberg, a case involving a defendant who drove his truck over the plaintiff, who was passed out in a field of tall grass. The Rikstad court rejected the appellate court’s approach, which considered the foreseeability of harm under the element of proximate causation. The court instead looked to the foreseeability of the harm to define the scope of the defendant’s duty to the plaintiff. The Rikstad court held that, for there to be a duty, the harm that occurs does not have to be one that was specifically anticipated, but must only have been within the general field of dangers that could be foreseen by a reasonable person. Applying this approach, the court found that the risk of the plaintiff being run over while passed out in the field was within the general field of dangers covered by the defendant’s duty to use reasonable care while driving a vehicle.

The general field of danger encompasses a broad range of harms; harms fall outside this scope only if “by no flight of the imagination” could they be foreseen. It does not matter that the harm was bizarre or that it came about as a result of a highly improbable sequence of events. It is also inconsequential that the harm could have happened earlier, but did not. For example, in Palin v. General Construction Co., the defendant argued that because oil could have previously been released from unsecured valves on a storage tank, but had never been before, the eventual release of the oil was not foreseeable. The court rejected this argument, holding that the fact that the oil spill could have occurred earlier did not mean that it was not foreseeable that the oil

102. See id. at 268, 456 P.2d at 357.
103. See id. at 268, 456 P.2d at 357–58.
104. See id.; see also Hartley v. State, 103 Wash. 2d 768, 798, 698 P.2d 77, 83–84 (1985) (discussing how the question of foreseeability is properly addressed under the element of duty, but is often mistakenly addressed under the element of proximate cause).
105. See Rikstad, 76 Wash. 2d at 268, 456 P.2d at 357–58.
106. See id.
107. See McLeod v. Grant County Sch. Dist., 42 Wash. 2d 316, 322, 255 P.2d 360, 364 (1953) (using example of when a neighbor shooting a baby sitter was found unforeseeable because it was most unusual, and the employer was therefore found not to have had a duty to protect against this harm).
108. See id.
110. See id. at 250, 287 P.2d at 327.
Electronic Court Records & Malpractice

would spill eventually. There is no firm rule determining the scope of the field of danger; rather, the field’s outer limits depend on the facts of each case.

B. There Must Be a Causal Link Between the Attorney’s Breach of a Specific Duty and the Client’s Harms

Establishing the causal link between the attorney’s breach of duty and the harm suffered by the client is often the most difficult part of a legal malpractice claim. To be held liable for the client’s harm, the attorney must be the proximate cause of the harm. Proving that the attorney was the proximate cause requires that the client show that the attorney was both the cause in fact of the harm, and the legal cause of the harm.

1. A Plaintiff Proves the Factual Link by Satisfying the “But For” Test

In Washington legal malpractice cases, the cause in fact test is satisfied if the harm would not have happened “but for” the attorney’s negligence. Under this standard, the client must prove that the attorney’s act “more likely than not” caused the client’s harm through a direct, unbroken sequence of events that would not have occurred without the attorney’s act. Cause in fact is generally for the trier of fact to decide, based upon the court’s instructions on how the “but for” standard is to be applied to the facts.

111. See id.
114. See Daugert, 104 Wash. 2d at 257, 704 P.2d at 603; Lord, supra note 112, at 1480. Cause in fact is generally a question for the jury unless only one reasonable result is possible. See Bullard v. Bailey, 91 Wash. App. 750, 757, 959 P.2d 1122, 1126 (1998).
116. See Daugert, 104 Wash. 2d at 257–58, 704 P.2d at 603; Hartley v. State, 103 Wash. 2d 768, 778, 698 P.2d 77, 83 (1985) (stating that “cause in fact is generally left to the jury” and that “such questions of fact are not appropriately determined on summary judgment unless but one reasonable conclusion is possible”).
2. Legal Causation Depends Primarily on Judicial Policy and Whether Defendant’s Act Actually Increased Plaintiff’s Risk of Harm

Legal causation represents the determination that, as a matter of law, the defendant’s action was a cause of the plaintiff’s injury. The test for proving that the attorney was the legal cause of the client’s harm is more ambiguous than the test for proving that the attorney was the cause in fact of the client’s harm. When deciding whether the attorney was the legal cause of the client’s harm, courts examine the facts of the case to determine “whether, as a matter of policy, the connection between the ultimate result and the act of the defendant is too remote or insubstantial to impose liability.” As part of making this policy determination, courts ask two questions: (1) did the defendant’s act increase the risk of the harm occurring, rather than triggering the harm by mere chance; and (2) was the harm caused by a substantial and unforeseeable intervening action. In addition to these two questions, courts must also consider any other relevant factors affecting whether, as a matter of policy, liability should follow from the defendant’s action. The Washington Pattern Jury Instructions’ definition of proximate causation requires that the plaintiff show that his injury occurred “in a direct sequence, unbroken by any new independent cause.” The jury decides the question of legal causation as a component of the overall element of proximate causation. Nevertheless, the element of legal causation is

---


118. Almost all prior legal malpractice cases have focused on determining whether the attorney’s action during trial was the cause in fact of an unfavorable outcome during the trial. There is only one legal malpractice case clearly focusing on the issue of legal causation. See Bullard, 91 Wash. App. at 759, 959 P.2d at 1127.

119. Kim, 143 Wash. 2d at 204, 15 P.3d at 1289 (quoting Tyner v. Dep’t of Soc. & Health Servs., 141 Wash. 2d 68, 82, 1 P.3d 1148, 1155–56 (2000) (internal citation omitted)); see City of Seattle v. Blume, 134 Wash. 2d 243, 252, 947 P.2d 223, 227 (1997). Legal causation is ultimately a question of law for the court to decide, see Kim, 143 Wash. 2d at 203, 15 P.3d at 1289, though the jury may be called upon to make factual determinations on issues such as foreseeability, see Greenleaf v. Puget Sound Bridge & Dredging Co., 58 Wash. 2d 647, 654, 364 P.2d 796, 802 (1961).

120. See DeWOLF & ALLEN, supra note 78, § 4.21, at 115.

121. See King v. City of Seattle, 84 Wash. 2d 239, 250, 525 P.2d 228, 234 (1974).

122. 6 WASH. SUPREME COURT COMM. ON JURY INSTRUCTIONS, WASHINGTON PRACTICE, WASHINGTON PATTERN JURY INSTRUCTIONS—CIVIL § 15.01, at 179 (2002); see also id. § 15.04, at 189, § 15.05, at 191.

123. See Hartley v. State, 103 Wash. 2d 768, 778, 698 P.2d 77, 83–84 (1985) (recognizing that while cause in fact is entirely a jury issue, the jury’s role in determining legal causation is limited to
generally more of an obstacle for plaintiffs in the realm of summary judgment or motions for judgment not withstanding the verdict, where the court will be called upon to make a specific determination regarding legal causation.124

*Bullard v. Bailey*125 is the only legal malpractice case in Washington that focuses on the question of whether the defendant’s action increased the risk of the harm occurring to the plaintiff.126 In that case, Bailey, the defendant-attorney, had an unlicensed business associate who held himself out as an attorney and eventually stole the settlement funds belonging to Bailey’s client.127 Bailey failed to take steps to correct his associate’s representations.128 Division One of the Washington State Court of Appeals held that Bailey’s failure to act increased the risk that his client would lose his settlement funds.129 Because the unlicensed associate was having financial difficulties, the court decided that it was not mere chance that the client lost his settlement funds.130 Therefore, the court held that, as a matter of policy, it was proper to hold Bailey liable for his client’s loss.131

While no Washington court has addressed the issue of intervening causation in a legal malpractice case, many courts have addressed it in general negligence cases. Under Washington tort law, an intervening act can persuade the court to find that the defendant’s act was not the legal cause of the plaintiff’s harm if the intervening act was unforeseeable.132

---

124. See Hartley, 103 Wash. 2d at 779, 698 P.2d at 83 (in reviewing motion for summary judgment the court distinguished between cause in fact as a primarily jury question and legal causation that requires “a determination of whether liability should attach as a matter of law given the existence of cause in fact”); King, 84 Wash. 2d at 250, 525 P.2d at 234 (1974) (stating that the “court still must adduce from the record whether, as a policy of law, legal liability should attach to the defendant if the other factual elements are proven”).


126. While it is the only case discussing the issue of legal causation, the opinion heavily intermingled its discussion on the scope of the duty, the defendant’s responsibility for increasing the risk of harm, and the question of whether there was a superseding cause. See, e.g., id. at 755–59, 959 P.2d at 1125–27. As a result, the case is better suited for illustrative uses than actual legal authority on any of these issues.

127. See id. at 759, 959 P.2d at 1124–25.

128. See id.

129. See id.

130. See id.

131. See id.

132. See id. at 758–59, 959 P.2d at 1127; DEWOLF & ALLEN, supra note 78, § 4.23, at 117.
An intervening act is foreseeable and will not sever liability if it was one of the hazards that made the defendant’s act negligent. An intervening act can be negligent or criminal. However, Washington courts have held that the negligence or criminality of the intervening act is not itself determinative, but only one factor to be considered in deciding whether the act was foreseeable.

In McLeod v. Grant County School District, the Washington State Supreme Court held that intervening acts are unforeseeable only if their occurrence is “so highly extraordinary or improbable as to be wholly beyond the range of expectability.” In McLeod, the school negligently left a darkened room under bleachers in the school’s gymnasium unlocked. During gym class, three male students took one of their female classmates into this room and sexually assaulted her. The school district argued that the harm suffered by the girl was not a foreseeable result of the school leaving the darkened room unlocked. The district maintained that the sexual assault should break the causal chain between the school’s negligence and the female student’s harm because it was a crime that could not reasonably have been foreseen. Applying this standard, the court decided that the male students’ sexual assault was foreseeable because the district could have foreseen the increased likelihood that some act of indecency would take place in such a location. Thus, the sexual assault did not prevent the school’s negligent act from being a legal cause of the girl’s harm.

133. See Palin v. Gen. Constr. Co., 47 Wash. 2d 246, 250, 287 P.2d 325, 327 (1955) (noting that “an intervening criminal act may be found to be reasonably foreseeable and, if so, liability may be predicated thereon,” and holding that oil tank owner could recover against construction company for lost oil despite intervening act of someone else opening tank release valve).
135. 42 Wash. 2d 316, 255 P.2d 360 (1953).
136. See id. at 323, 255 P.2d at 364 (citing Berglund v. Spokane County, 4 Wash. 2d 309, 103 P.2d 355 (1940)).
137. See id. at 317, 255 P.2d at 361.
138. See id. at 318, 255 P.2d at 361.
139. See id. at 319, 255 P.2d at 362.
140. See id.
142. See id.
Electronic Court Records & Malpractice

IV. ATTORNEYS WHO FAIL TO REDACT IN COMPLIANCE WITH THE POLICY SHOULD BE LIABLE FOR LEGAL MALPRACTICE IF CLIENTS' IDENTITIES ARE STOLEN

The Policy’s redaction provision is a substantial step towards protecting clients’ identities, but it may only truly be effective if it is backed by the possibility of a cause of action against those who fail to abide by it. The ability to pursue a legal malpractice action would provide this effect. Clients face three hurdles to bringing this kind of malpractice claim successfully against their attorneys. First, they must convince courts to accept the Policy as evidence of a specific duty that attorneys owe to clients to redact sensitive information.143 Second, they must further show that this duty encompasses the harms resulting from identity theft.144 Finally, clients must establish that their attorneys’ breach of the duty to redact proximately caused the harms resulting from the loss of their identities.145 Under existing Washington malpractice precedent, clients should be able to overcome all three obstacles.

A. A Malpractice Action Is Necessary Because the Policy Does Not Provide an Enforcement Mechanism

The federal courts’ rapid implementation of electronic access to court records endangers clients by exposing litigants’ personal information to identity theft.146 Identity thieves can use this information to steal clients’ identities, plunder their financial accounts, and rack up bad debt in clients’ names.147 Today, the Policy is the only protection that clients have against identity theft.148 Yet, standing alone, it insufficiently safeguards clients’ identities. At best, the Policy merely reminds courts and attorneys to look after litigants’ interests before making sensitive information available electronically.149 Unfortunately, the Policy does not expressly protect clients once their personal information has already been stolen. It may be difficult, if not highly unlikely, for clients to collect from the thieves who stole their identities given the transitory

143. See supra notes 87–99 and accompanying text.
144. See supra notes 100–11 and accompanying text.
145. See supra notes 112–42 and accompanying text.
146. See supra note 6 and accompanying text.
147. See supra notes 40–49 and accompanying text.
148. See supra notes 5–9 and accompanying text.
149. See supra notes 55–73 and accompanying text.
nature and economic recourses of identity thieves. Therefore, to make themselves whole, clients should sue their attorneys for malpractice when these attorneys fail to redact personal information from documents before filing those documents with the court.

B. **The Policy Is Evidence That Attorneys Have a Duty To Redact Their Clients’ Personal Information Before Filing Documents with the Court**

The Policy requires that attorneys redact their clients’ personal information from documents before filing those documents with the court. Thus, the Policy provides evidence that redaction is a specific duty that attorneys owe their clients. Whether Washington courts will accept the Policy into evidence depends on their interpretation of the reach of the Washington State Supreme Court’s decision in *Hizey v. Carpenter*. The *Hizey* decision prohibits clients from using the RPCs as evidence of their attorneys’ specific duties, and could also be read to prohibit the use of the Policy as evidence of the duty to redact. Both the RPCs and the Policy were products of judicial policymaking rather than legislative enactment, and Washington courts might be asked to focus on the origin of the Policy rather than the specific rationales that supported the outcome in *Hizey*. However, if Washington courts look beyond the origin of the documents, they will see that the three justifications supporting the *Hizey* court’s exclusion of the RPCs do not support exclusion of the Policy. After evaluating the rationales provided in *Hizey*, Washington courts should allow clients to submit the Policy as evidence of attorneys’ specific duty to redact clients’ personal information from documents before filing those documents with the court.

1. **The Policy Leaves Open the Possibility of Its Use To Support Existing Causes of Action**

The *Hizey* court’s first rationale for rejecting the RPCs as proof of attorneys’ specific duties was that the RPCs expressly deny establishing any specific “standards for civil liability of lawyers for professional

---

150. See *supra* note 6.
151. See *supra* notes 68–73 and accompanying text.
Electronic Court Records & Malpractice

conduct.” The Policy, on the other hand, does not expressly preclude clients from using it to demonstrate their attorneys’ specific duties. Because a cause of action for legal malpractice already exists, clients will only use the Policy to prove a specific duty owed to them, not to create a private cause of action. The Policy does not expressly preclude clients from such a use. Therefore, the first *Hizey* factor does not bar use of the Policy as evidence of the attorneys’ duty.

2. **The Policy’s Redaction Provision Protects Only Clients, Not Courts**

   The second reason the *Hizey* court gave for rejecting admission of the RPCs was that they reflect mixed, and at times conflicting, public and private obligations for attorneys. The court found that the RPCs, as adopted by the Washington State Supreme Court, were designed primarily to protect the integrity of the judicial system, with the rights of individual clients being of lesser concern. The *Hizey* court feared that allowing the RPCs into evidence would encourage attorneys to interpret the RPCs primarily as protecting clients, minimizing the role that the RPCs play in maintaining the integrity of the court system. While it is true that the Policy, like the RPCs, protects the integrity of the judicial system, it does so by calling on the courts to impose Rule 11 sanctions against attorneys who fail to manage their court filings carefully. The Rule 11 provision is a distinct part of the Policy, and unrelated to the duty of attorneys to redact their own clients’ information. This type of distinction could not be as readily made between the RPCs’ court protections and client protections, since as the *Hizey* court noted, all of the RPCs’ provisions are designed to remind attorneys “that their first loyalty is to the court.” Because the Policy’s protections for the system and for the individual do not conflict with each other, the *Hizey* decision should not bar use of the Policy as evidence of attorneys’ duty to redact.

153. See *Hizey*, 119 Wash. 2d at 258, 830 P.2d at 650; *supra* note 96 and accompanying text.
154. See *Hizey*, 119 Wash. 2d at 263, 830 P.2d at 653; text accompanying *supra* note 97.
155. See *Hizey*, 119 Wash. 2d at 263, 830 P.2d at 653; text accompanying *supra* note 97.
156. See *Hizey*, 119 Wash. 2d at 263, 830 P.2d at 652–53.
157. See *supra* note 63 and accompanying text.
3. **The Policy’s Mechanical Duty To Redact Is Well Suited to Being Used as Evidence of a Specific Duty**

Finally, the *Hizey* court raised concerns about the fact that the RPCs describe only general obligations without stating particular duties owed to clients.\(^{159}\) The court emphasized that the RPCs represent a vague scheme for evaluating attorneys’ ethical responsibilities to both their clients and the court.\(^{160}\) For example, the court pointed to several sections of the RPCs that require attorneys to avoid conflicts of interest without specifying how.\(^{161}\) The court rejected the RPCs’ ethical scheme as too vague to establish duties owed to clients in particular cases.\(^{162}\) In contrast, the Policy calls on attorneys to take a specific action—redaction of four types of personal information—in order to protect their clients.\(^{163}\) Unlike the RPCs, there is much less subjectivity involved in the Policy’s requirement to redact specifically defined types of information. Therefore, under *Hizey*, the Policy is appropriate evidence of a specific duty.

**C. Identity Theft Is a Foreseeable Harm Within the General Field of Danger Contemplated by the Duty To Redact**

In order to prove that the attorney committed malpractice, the plaintiff-client must establish not only that the attorney breached a specific duty, but also that the harm resulting from that breach was within the general field of danger contemplated by the duty.\(^{164}\) Clients should be able to demonstrate that the harms resulting from identity theft are within the general field of danger that the duty to redact was designed to protect against. For a harm to be within the general field of danger, it must be foreseeable,\(^{165}\) such as the client’s loss of funds in *Bullard*,\(^{166}\) or the injuries to the plaintiff sleeping in the long grass in *Rikstad*.\(^{167}\)

---

159. See id.; text accompanying supra note 98.
160. See *Hizey*, 119 Wash. 2d at 263, 830 P.2d at 652; text accompanying supra note 98.
161. See *Hizey*, 119 Wash. 2d at 258, 830 P.2d at 650.
162. See supra note 98 and accompanying text.
163. See supra note 69 and accompanying text.
164. See supra notes 100–11.
165. See supra notes 100–11.
166. See supra notes 125–31 and accompanying text.
167. See supra notes 102–06 and accompanying text.
The widespread and public discussion on the risk of identity theft should include the harms resulting from identity theft in the realm of events within the “flight of the imagination” of a reasonable attorney. The harm resulting from loss of identity is foreseeable for three reasons. First, identity theft is a major crime that has received widespread attention. Second, critics of the Policy have already predicted that thieves would use information from court records to steal clients’ identities. Third, the Policy recognizes that attorneys will have to protect their clients from the risks associated with making court filings available electronically, including the risk that identity thieves will comb those filings for clients’ personal information.

Nevertheless, some defendants might argue that the harms resulting from identity theft are not foreseeable because there have been few problems with electronic access thus far. This argument is not likely to succeed, however, because Washington precedent holds that the fact that a specific harm could have occurred earlier, but did not, is insufficient to make the eventual harm unforeseeable. Therefore, clients should be able to establish that harms resulting from identity theft are within the general field of danger encompassed by the duty to redact.

D. Clients Should Be Able To Prove That the Failure To Redact Was the Proximate Cause of the Harm Suffered as a Result of Identity Loss

Once clients demonstrate that their attorneys owed them a duty to redact their personal information from court filings, and that the duty to redact protects against the harms resulting from identity theft, they must still prove that their attorneys’ negligence caused them to suffer those harms. The element of causation has long been recognized as the most formidable obstacle in a legal malpractice claim because even the most

---

168. See McLeod v. Grant County Sch. Dist., 42 Wash. 2d 316, 322, 255 P.2d 360, 364 (1953); text accompanying supra note 107.
169. See supra notes 42–49 and accompanying text.
170. See supra note 41.
171. See EHLERS ET AL., supra note 6, at § 3.13; Givens, supra note 6.
172. See EHLERS ET AL., supra note 6, at § 3.13; Givens, supra note 6.
173. See supra note 53 and accompanying text.
174. See Palin v. Gen. Constr. Co., 47 Wash. 2d 246, 250, 287 P.2d 325, 327 (1955) (holding that the fact that the oil spill could have occurred earlier did not mean that it was not foreseeable that the oil would spill eventually).
175. See supra notes 112–18 and accompanying text.
egregious breach of duty has no legal consequence without proof of causation. To establish causation, clients must first show that their attorneys’ actions were the cause in fact of their harm. Then, they must prove that their attorneys’ actions were the legal cause of their harm.

1. The Simpler Test for Cause in Fact

To establish that their attorneys were the cause in fact of their harm, clients must prove that their attorneys’ failure to redact triggered a direct and unbroken sequence of events that would not have occurred but for the attorneys’ negligence. It is likely that triers of fact will find that attorneys who fail to redact are the cause in fact of the harms resulting from identity loss when the thief used information contained in court records. Clients will likely need to obtain information tying the thief to the PACER system in order to show that their loss occurred because of the opportunity created by the attorneys’ failure to redact, rather than through some other means for the thief to obtain their information.

Proving that the thief obtained the client’s information from court records should be assisted by the PACER system’s ability to track who has viewed specific filings.

In sum, because stealing identities is a crime of opportunity, a jury could reasonably view an attorney’s failure to redact as creating the opportunity for an identity thief to assume the client’s identity. Without the attorney failing to redact, there would be no chain of events leading to the client suffering harm from a lost identity. Thus, the attorney’s failure to redact is a cause in fact of the client’s harm.

176. See Lord, supra note 112, at 1480.
177. See supra notes 114–15 and accompanying text.
178. See supra notes 117–24 and accompanying text.
180. This Comment assumes the client will be able to demonstrate that the thief obtained the personal information through the unredacted court records. Although this may be a difficult task, the PACER system has been set up so that the use of a login identity will allow courts to maintain “an electronic trail which can be retraced in order to determine who accessed certain information if a problem arises.” See POLICY, supra note 3.
181. See id.
182. See Givens, supra note 6.
2. **Establishing Legal Causation: Failing To Redact Increases the Risk of Identity Theft**

In order to prove that their attorneys were the legal cause of the harms resulting from their lost identity, clients must show that their attorneys’ actions increased the risk of the harms happening, and that those harms were not the result of an intervening act. Proving that an attorney’s action increased the risk of identity theft should not be difficult. In *Bullard*, the court held that the attorney’s failure to supervise his associate increased the client’s risk of losing his settlement money because the associate was having financial difficulties and should not have been given the opportunity to steal the client’s funds. Because identity theft is also a crime of opportunity, attorneys who fail to redact their clients’ personal information from court filings similarly increase the risk of harm to their clients by giving thieves the chance to steal their clients’ identities. Furthermore, the existence of the special redaction protection to prevent identity theft should weigh heavily in favor of finding that an attorney’s failure to abide by the protective measure resulted in an increased risk of the harm it was meant to protect against. That an identity thief might use PACER records to obtain information was the type of concern that prompted the creation of the Policy and the impetus behind the requirement that a password be obtained before accessing PACER is allowed. Because the Judicial Conference Committee created the redaction protection in part to prevent court records from assisting in crimes like identity theft, when identity theft based on failure to redact occurs, it would be difficult to argue that the theft occurred by mere chance.

Once clients establish that their attorneys’ failure to redact increased their risk of suffering harm, they must be prepared to respond to the argument that identity thieves are intervening actors who break the causal chain between the attorneys’ negligent failure to redact and the clients’ harms. An intervening act, even if criminal, will only sever

---

183. See supra notes 117–42 and accompanying text.
185. See Givens, supra note 6.
186. See POLICY, supra note 3; text accompanying supra note 36.
187. See also FAQ, supra note 8, at 2.
188. See supra notes 132–34 and accompanying text.
legal causation if it was unforeseeable. To be unforeseeable, the intervening act must be “so improbable or exceptional as to be wholly unexpected.”

As shown above, the harms resulting from identity theft are foreseeable. Further, because the act of identity theft is the only way for the harms resulting from identity theft to occur, the act of identity theft is also foreseeable. The intervening act’s status as a crime only relates to whether it was foreseeable, and because this crime is not wholly unexpected, its criminal nature is not relevant. Therefore, any argument that identity theft is an unforeseeable intervening act should fail.

Furthermore, holding attorneys liable for the harms resulting from identity theft is sound judicial policy based on the catch-all rule allowing courts to look at any other relevant factors. First, a legal malpractice claim is likely the only meaningful recourse available to clients who lose their identities due to their attorneys’ negligence. Second, the threat of malpractice liability may be an effective means of proactively inducing attorneys to protect their clients’ interest when filing court materials. In addition, a major implication of failing to find foreseeability would be that, despite vocal concerns prior to cases of identity theft actually occurring, the first victims of identity theft would be forced to bear their losses simply because they were first. Finally, in terms of efficient risk management, the attorney as the skilled professional is likely the best party to bear the burden of ensuring that personal information is omitted from documents filed by the attorney. Given the potentially severe impact from denying legal cause as a matter of policy, combined with the attorney’s status as the better risk avoider, strong policy justifications exist for recognizing legal cause in failure to redact cases.

189. See supra notes 133–34, 136 and accompanying text.
190. See supra note 136 and accompanying text.
191. See supra note 134 and accompanying text.
192. See supra note 134 and accompanying text.
193. See supra note 121 and accompanying text.
194. See supra note 6; Givens, supra note 6. Furthermore, identities are inanimate objects and do not steal themselves.
195. See supra note 134 and accompanying text.
196. See supra note 136 and accompanying text.
197. See supra note 121 and accompanying text.
198. See supra note 6; Givens, supra note 6.
199. See Scott Peterson, Comment, Extending Legal Malpractice Liability to Nonclients—The Washington Supreme Court Considers the Privity Requirement—Bowman v. John Doe Two, 61 WASH. L. REV. 761, 768 (1986) (noting that a purpose of negligence law is to deter culpable behavior, and applying this concept to legal malpractice); see also Harbeson v. Parke-Davis, Inc., 98 Wash. 2d 460, 481, 656 P.2d 483, 496 (1983) (recognizing the deterrent purpose of negligence law in the context of medical malpractice).
Electronic Court Records & Malpractice

V. CONCLUSION

The Policy attempts to provide users of the judicial system substantial benefits through electronic access to court records. By reasonably interpreting existing legal malpractice precedent, Washington courts can ensure that these benefits are not achieved at clients’ expense. Washington clients can be effectively protected under the Policy as long as the Washington courts are willing to adapt legal malpractice law to the new reality of electronic access to court records. A reasonable development of existing legal malpractice law can help to ensure that clients have some effective recourse when their attorneys fail to adequately protect them. Washington courts should admit the Policy as evidence of an attorney’s specific duty of care and recognize proximate causation in legal malpractice cases stemming from an attorney’s failure to redact as required by the Policy.
BALANCING CONSUMER INTERESTS IN A DIGITAL AGE: A NEW APPROACH TO REGULATING THE UNAUTHORIZED PRACTICE OF LAW

Cristina L. Underwood

Abstract: States have traditionally relied on unauthorized practice of law statutes and court rules to restrict nonlawyers from providing legal services. A majority of courts assess compliance with these statutes by applying set practice of law definitions and restrictive court precedent to nonlawyer activity. These methods of enforcement have failed to balance consumer protection concerns with the public’s need for access to affordable legal services. Most state practice of law definitions have proven inflexible, broadly barring the practice of law by nonlawyers, with few exceptions. Courts interpreting unauthorized practice statutes have created bright-line rules that favor consumer protection, failing to incorporate additional factors such as the reliability of the nonlawyer service. These approaches are ill-equipped to adapt to changing technologies, such as the development of interactive software and online programs that enable nonlawyers to provide document preparation services to specific consumers. States should incorporate a test that requires the assessment of four factors significant to legal access and consumer protection concerns into their unauthorized practice of law statutes or court rules. Such a test would assist courts in determining whether a particular nonlawyer service is beneficial to consumers, and would provide the flexibility necessary for states to adapt more readily to new technologies.

Erin and her husband would like to obtain an uncontested divorce. \(^1\) The couple considers hiring an attorney, but they cannot afford the $228 hourly rate. \(^2\) Erin has heard that uncontested divorces are relatively simple, particularly when there are no children involved. She believes she can submit the legal forms herself, but would like the assistance of a document preparation service, which would select and prepare the appropriate forms for her. While exploring several nonlawyer Internet services, Erin finds LawDocs.com. \(^3\) LawDocs.com prepares uncontested divorce papers online. The service consists of a two-step process. First, a consumer must submit responses to specific questions. Trained nonlawyers at LawDocs.com then use the information to prepare the customized forms, delivering the documents to the consumer within two

---

1. Hypothetical created by author.
2. The average hourly rate for lawyers in the United States is $228. See Jennifer Mulrean, Get an Online Divorce, MSN Money, at http://moneycentral.msn.com/content/CollegeandFamily/Suddenlysingle/P56205.asp (last visited Jan. 11, 2004) (citing SURVEY OF LAW FIRM ECONOMICS (2002)).
3. Fictional company and website created by author.
days. For Erin and her husband the total cost of the service would be approximately $250. The couple is interested in LawDocs.com, which would enable them to obtain the divorce at a more affordable price, but they are concerned about the service’s reliability and confidentiality.

The experience of Erin and her husband illustrates some of the benefits and concerns associated with nonlawyer legal services. While nonlawyers may provide consumers with economical alternatives to legal representation, their services also raise consumer protection concerns. For example, the couple’s knowledge of LawDocs.com’s reputation would be restricted to the information provided by the company’s website and the Better Business Bureau. Erin and her husband would not know whether LawDocs.com has misrepresented its level of experience or reliability. Unlike lawyers, nonlawyer services generally are not regulated by state bar associations, are not required to acquire specific levels of training, and are less likely to be disciplined for poor performance. In addition, because the state legal rules of professional conduct apply only to lawyers, information submitted to a nonlawyer does not enjoy the same protection against conflicts of interest and breaches of confidentiality.

States generally regulate the unauthorized practice of law through statutes and court rules that restrict the types of services nonlawyers can provide. States take one of two main approaches to the regulation of the unauthorized practice of law: (1) the development and enforcement of a practice of law definition, created by statute or court rule, or (2) the establishment, by a state court, of bright-line rules for what is not the practice of law. These approaches lack the flexibility necessary to strike a proper balance between legal access and consumer protection concerns. Moreover, these approaches are incapable of keeping pace with the development of interactive software and online technology.

Without attempting to define the practice of law, this Comment proposes that in the interest of protecting consumer interests, state

7. See infra Part II.B–C.
8. See infra Part II.B–C.
9. See infra Part II.B–C.
Consumer Interests & the Unauthorized Practice of Law

lawmakers should adopt a test with four factors for courts to apply when determining whether a given activity is the unauthorized practice of law. This determination would take place within the framework of a general unauthorized practice of law statute or practice of law definition. The adoption of this proposed test would enable states to retain general practice of law definitions capable of adapting to innovative products and beneficial nonlawyer services. Part I of this Comment discusses the development of the unauthorized practice of law and the consumer protection and access to justice considerations raised by nonlawyer activities. Part II assesses current methods of defining and regulating the practice of law. In response to the shortcomings of existing approaches, Part III argues that state lawmakers should adopt a test that requires courts to weigh legal access and consumer protection concerns when deciding whether nonlawyers can engage in the practice of law.

I. PROSCRIPTIONS ON THE UNAUTHORIZED PRACTICE OF LAW ATTEMPT SIMULTANEOUSLY TO PROTECT CONSUMERS AND PROVIDE ACCESS TO JUSTICE

States have regulated the practice of law, through statutes and court rules, since the founding of the legal profession in America. Traditionally, states have employed unauthorized practice regulations as a means of protecting both consumers and the interests of the legal profession. With the unmet legal needs of consumers on the rise, however, members of the legal profession are beginning to recognize that these regulations must strike a balance between protecting consumers and protecting access to legal resources.

A. Regulation of the Unauthorized Practice of Law Has Been Motivated by Both Consumer Protection Concerns and the Private Interests of the Legal Profession

A combination of public interest and personal interest considerations has motivated states to implement unauthorized practice restrictions.
The American Bar Association (ABA) supports practice of law restrictions as necessary to protect consumers from excessive risks of harm. Members of the legal profession have cited various problems in allowing nonlawyers to practice law. First, nonlawyers sometimes advise clients on complex legal issues without understanding the intricacies involved. Erroneous advice can seriously harm a client’s case and lead to further expenses and litigation. Second, clients may have false expectations as to the expertise of the nonlawyer or the scope of the service provided. Consumers expect legal service providers to have legal training or experience and often anticipate that nonlawyers will effectively assess their legal needs.

Finally, clients of nonlawyers practicing law forgo the protection afforded by state rules of professional conduct regarding conflicts of interest and confidentiality, and by attorney-client privileges. Conflicts of interest routinely emerge when nonlawyers represent clients with directly adverse interests or when nonlawyers stand to benefit from the sale of a legal instrument that may not be appropriate for a particular consumer. For example, an insurance agent’s objective in creating an estate plan for a customer may be the sale of an insurance policy rather than the customer’s best interests. Consumers using services provided by nonlawyers are also more vulnerable to breaches of confidentiality, particularly when nonlawyers fail to adhere to comprehensive privacy policies or to take reasonable precautions to prevent the unintentional dissemination of confidential information. In addition, due to the

15. See AM. BAR ASS’N COMM’N ON NONLAWYER PRACTICE, supra note 13, at 126–27.
16. See Tebo, supra note 4, at 41–42.
17. See id.
18. See id. at 40.
19. See id.
20. See RESTATEMENT (THIRD) OF LAW GOVERNING LAWYERS § 4 (2000); see also id. § 70 (stating that “[p]rivileged persons . . . are the client . . . the client’s lawyer, agents of either who facilitate communications between them, and agents of the lawyer who facilitate the representation”).
21. See MODEL RULES OF PROF’L CONDUCT R. 1.7 (2002) (noting that a conflict of interest is created when lawyers represent clients with directly adverse interests).
22. See In re Mid-Am. Living Trust Assocs., Inc., 927 S.W.2d 855, 860 (Mo. 1996) (noting that “a conflict of interest exists between those who benefit from the sale of a particular legal instrument and the client for whom that legal instrument may not be appropriate”).
23. See Christensen, supra note 6, at 206.
25. See MODEL RULES OF PROF’L CONDUCT R. 1.6 cmt. (requiring lawyers to take “reasonable precautions to prevent . . . information from coming into the hands of unintended recipients”).
Consumer Interests & the Unauthorized Practice of Law

absence of an attorney-client privilege, courts may compel nonlawyers to testify about communications involving their clients. Nonlawyers generally seek to avoid some of these problems by providing disclaimers that inform consumers of potential conflicts of interest or the absence of an attorney-client privilege. These disclaimers, however, may lack important information and are often ignored by consumers.

Despite the legal profession’s concerns about nonlawyer services and products, some commentators argue that there is relatively little evidence to suggest that such activities pose a high risk to consumers. Injured consumers do not initiate many court cases involving the unauthorized practice of law; rather, state bar committees usually prosecute these cases after conducting an independent committee investigation. Neither the secondary legal literature nor court decisions suggest that the public perceives unauthorized practice as a substantial danger. In addition, consumer groups and governmental agencies have criticized certain practice of law restrictions as anti-competitive and costly to consumers. Some contend that a primary motivating factor behind

26. Communications between nonlawyers and clients are generally not privileged unless the communication is made to a person “who is a lawyer or who the client or prospective client reasonably believes to be a lawyer.” See Restatement (Third) of the Law Governing Lawyers §§ 70, 72. However, in rare instances courts have extended the attorney-client privilege to nonlawyers who are “lawfully performing the functions of an attorney, such as representing a client in an administrative proceeding.” See Christopher B. Mueller & Laird C. Kirkpatrick, Federal Evidence § 183 (2d ed. 1994).


28. See Tebo, supra note 4, at 40.

29. See id.


31. See Christensen, supra note 6, at 203. States generally enforce the unauthorized practice of law through either court orders or the advisory rulings of bar committees. Penalties may range from civil fines and injunctions to criminal fines and imprisonment. Enforcement, however, has been inconsistent, with several states not actively enforcing their unlawful practice of law regulations. Some states have indicated that insufficient funding frustrates enforcement efforts. See AM. BAR ASS’N COMM’N ON NONLAWYER PRACTICE, supra note 13, at 23, 119; see also AM. BAR ASS’N STANDING COMM. ON CLIENT PROT., AM. BAR ASS’N CTR. FOR PROF’L RESPONSIBILITY, 1999 SURVEY OF UNAUTHORIZED PRACTICE OF LAW COMMITTEES 1–2 (1999).

32. See Christensen, supra note 6, at 201; see also Rhode, supra note 30, at 43.

unlawful practice regulations is self-interest on the part of the legal profession.\textsuperscript{34}

\section*{B. The Legal Profession Has Recognized a Need To Protect Public Access to Nonlawyer Services}

The legal profession’s interest in protecting consumers by restricting nonlawyer activity is inconsistent with its interest in enhancing public access to legal services.\textsuperscript{35} According to an ABA study, “as many as 70\% to 80\% or more of low-income persons are unable to obtain legal assistance even when they need and want it.”\textsuperscript{36} The cost of attorney services has been a leading factor in the growth of pro se litigation.\textsuperscript{37} A 1991 study by the National Center for State Courts found that only twenty-eight percent of divorces proceed with both parties represented by an attorney,\textsuperscript{38} and attorneys draft only half of all estate documents and document real estate transactions nationwide.\textsuperscript{39} Many low- and moderate-income households simply cannot afford the cost of personal legal services.\textsuperscript{40} Even households that can afford such services sometimes opt to resolve simple legal transactions and claims themselves, rather than paying an attorney.\textsuperscript{41}

Nonlawyers can provide unrepresented consumers with economical alternatives to lawyer services. Traditional nonlawyer services include self-help books and do-it-yourself kits, both of which garnered popularity and attracted judicial scrutiny in the 1960s and 1970s.\textsuperscript{42} Businesses also provide personalized nonlawyer legal services in many

\begin{itemize}
  \item \textsuperscript{34} See Ronald D. Rotunda, Legal Ethics—The Lawyer’s Deskbook on Professional Responsibility § 39-2 (2002–2003).
  \item \textsuperscript{35} See Am. Bar Ass’n Comm’n on Nonlawyer Practice, supra note 13, at 126.
  \item \textsuperscript{36} See id. at 77.
  \item \textsuperscript{37} William Hornsby, Improving the Delivery of Affordable Legal Services Through the Internet: A Blueprint for the Shift to a Digital Paradigm (Nov. 1999), at http://www.unbundledlaw.org/program/program.htm.
  \item \textsuperscript{38} See Hornsby, supra note 37.
  \item \textsuperscript{39} See Tebo, supra note 4, at 40.
  \item \textsuperscript{40} See id.
  \item \textsuperscript{41} See id.
  \item \textsuperscript{42} Do-it-yourself kits provide consumers with legal forms and instructions on specific areas of law, such as family law, tax law, or wills. See Fla. Bar v. Stupica, 300 So. 2d 683, 684–85 (Fla. 1974); see also Or. State Bar v. Gilchrist, 538 P.2d 913, 914 (Or. 1975). For cases examining traditional nonlawyer services, see, for example, State v. Cramer, 249 N.W.2d 1, 2 (Mich. 1976) (divorce kits); N.Y. County Lawyers’ Ass’n v. Dacey, 283 N.Y.S.2d 984 (App. Div.), rev’d, 234 N.E.2d 459 (N.Y. 1967) (self-help books); Gilchrist, 538 P.2d at 913 (divorce kits).
\end{itemize}
Consumer Interests & the Unauthorized Practice of Law

states. These businesses include tax preparation services, debt collection agencies, title examination companies, accounting firms, trust departments of banking institutions, and real estate agencies.

The emergence of interactive computer technologies has spurred a new wave of products, contributing to the recent expansion of nonlawyer services. These products assist consumers in preparing legal documents in a variety of areas including business incorporation, living trusts, powers of attorney, uncontested divorce, and copyright applications. The programs take consumers through a systematic interview process, select the state-appropriate legal forms, and assemble the documents based on consumer information. Software programs, like Quicken® WillMaker, use decision-tree software to select the documents and input the consumer information. In contrast, online services, such as LegalZoom™, generally have nonlawyers select and complete the consumer forms via the Internet. These services enable nonlawyers to customize documents while avoiding traditional in-person consultations.

II. EXISTING METHODS FOR REGULATING THE UNAUTHORIZED PRACTICE OF LAW DISCOURAGE BROAD CONSUMER ACCESS TO INTERACTIVE Technologies

The unauthorized practice of law is the practice of law by a person, generally a nonlawyer, who lacks authorization to practice in a given jurisdiction. Some states regulate the unauthorized practice of law through the implementation of detailed definitions of the practice of law. Other states have shied away from creating a particularized definition of the practice of law, relying instead on courts’ interpretations of broad practice of law definitions and unauthorized practice of law regulations. Neither of these approaches, however, ensures access to interactive technologies.

43. See AM. BAR ASS’N COMM’N ON NONLAWYER PRACTICE, supra note 13, at 43.
44. See id. at 43–44.
46. See, e.g., sources cited supra note 45.
47. See Quicken® WillMaker Plus 2004, supra note 45.
48. See LegalZoom™, supra note 45.
49. See BLACK’S LAW DICTIONARY 1192 (7th ed. 1999).
50. See infra Part II.C.
A. The Difficulty in Crafting a Practice of Law Definition

Courts have found applying the unauthorized practice of law to specific nonlawyer activities to be a challenge because of the difficulty states have had defining the practice of law.\textsuperscript{51} States take two approaches to defining the practice of law. Some states provide only a general definition. For example, New Mexico defines the practice of law by court rule simply as:

(1) representation of parties before judicial or administrative bodies; (2) preparation of pleadings and other papers, incident to actions and special proceedings; (3) management of such actions and proceedings; and (4) noncourt-related activities, such as: (a) giving legal advice and counsel; (b) rendering a service which requires use of legal knowledge or skill; and (c) preparing instruments and contracts by which legal rights are secured.\textsuperscript{52}

A general practice of law definition does not preserve access to specific nonlawyer activities.

In contrast, other states define the practice of law and then take an additional step, offering specific exceptions that allow for certain nonlawyer activities. For example, Washington provides eleven specific exceptions to the practice of law, allowing nonlawyers to engage in certain activities, such as serving as courthouse facilitators and mediators, participating in labor negotiations, and selling legal forms.\textsuperscript{53}

Most states follow this second approach and authorize nonlawyers to perform limited legal services in specific areas.\textsuperscript{54} Depending on the state, these services may include drafting certain legal documents and attending (or in some cases participating) in administrative proceedings.\textsuperscript{55}

Despite state efforts to define the practice of law, several courts have noted that an all-encompassing practice of law definition is not feasible because "such practice must necessarily change with the ever-changing business and social order."\textsuperscript{56} Some members of the legal profession

\textsuperscript{51} See infra Part II.B.1.
\textsuperscript{52} N.M. R. ANN. 20-102.
\textsuperscript{53} See WASH. CT. GEN. R. 24.
\textsuperscript{54} See AM. BAR ASS’N STANDING COMM. ON CLIENT PROT., supra note 31, at 2.
\textsuperscript{55} See id.
\textsuperscript{56} See Fla. Bar v. Brumbaugh, 355 So. 2d 1186, 1191–92 ( Fla. 1978) (quoting State v. Cramer, 249 N.W.2d 1, 7 (Mich. 1976)); see also Iowa Supreme Court Comm. on Unauthorized Practice of Law v. Sturgeon, 635 N.W.2d 679, 685 (Iowa 2001); In re Campaign for Ratepayers’ Rights, 634
argue that a single definition cannot work in every situation because whether a person is engaged in the practice of law depends on context (i.e., if a client reasonably believes he or she has hired a lawyer, then the service provider is practicing law). The Federal Trade Commission (FTC) has also criticized state use of broad practice of law definitions for its failure to accommodate access to emerging technologies. According to the FTC, under a broad practice of law definition, “[i]nteractive web sites that help consumers write their own legal documents might be found to be practicing law, and Internet-based lenders would likely find that they could not complete real estate or loan closings without hiring a local attorney in the state where the property is located.”

Even with the limitations inherent in a set practice of law definition, the ABA and several states have recently shown a renewed interest in creating comprehensive practice of law definitions. According to the ABA, this movement has been in response to an increasing number of situations in which nonlawyers provide services that are difficult to classify as constituting the practice of law. Technological advancements in the provision of nonlawyer services have contributed to this ambiguity. The ABA believes practice of law definitions encourage consistent enforcement of unauthorized practice of law statutes and reduce the growing uncertainty associated with courts’ interpretations of the unauthorized practice of law.
B. Current Practice of Law Definitions Do Not Accommodate Interactive Technologies

State and ABA efforts to define the practice of law have been unsuccessful in creating a definition that balances consumer protection and legal access concerns. Only one state, Texas, has created a practice of law definition that preserves access to interactive technologies, but this access has come at the expense of consumer protection. Although the ABA encourages states to weigh competing consumer interests when crafting practice of law definitions, the ABA has not been able to provide an example of a practice of law definition that effectively balances these concerns.

1. Most State Practice of Law Definitions Do Not Provide Exceptions for Interactive Technologies

States generally favor consumer protection at the expense of legal access, classifying most nonlawyer activities as the practice of law. Several state definitions have narrow exceptions allowing nonlawyers to provide limited services, such as the preparation of legal documents by real estate agents or tax accountants. Three states—Washington, Arizona, and Texas—have adopted definitions that allow for greater nonlawyer involvement in the preparation and sale of legal documents. To varying degrees, these definitions place fewer restrictions on nonlawyer activity than the definitions followed by other states.

Washington’s practice of law definition, adopted by court rule in 2001, broadly defines the practice of law to include in-court representation; the giving of legal advice or counsel for a fee; and the selecting, drafting, or completing of legal documents or agreements that affect the legal rights of entities or persons. The definition then provides eleven exceptions to the rule, permitting certain nonlawyer

63. See infra notes 84–89 and accompanying text.
64. See infra notes 94–99 and accompanying text.
65. See infra notes 90–93 and accompanying text.
67. See id.
68. See infra notes 70–73, 77, 85 and accompanying text.
69. See WASH. CT. GEN. R. 24.
70. See id.
Consumer Interests & the Unauthorized Practice of Law

activities regardless of whether they constitute the practice of law. 71 The eighth exception provides for the “sale of legal forms in any format.”72 The eleventh exception is a catchall provision that allows for the practice of law by nonlawyers, when permitted by the Washington State Supreme Court.73

Because Washington’s practice of law definition is relatively new, its application to interactive technologies remains untested. For example, the eighth exception, allowing for the sale of legal forms, may not extend to interactive services, such as Quicken® WillMaker, which involve actual document preparation.74 In addition, the Washington State Supreme Court has not yet used the new deference accorded to it under the eleventh exception to expand the scope of the exceptions to include interactive technologies.75

Arizona’s practice of law definition attempts to balance consumer protection concerns with access to legal services by allowing only “certified” nonlawyers to prepare legal documents.76 Arizona defines the practice of law to include the preparation of “any document in any medium intended to affect or secure legal rights for a specific person or entity,” but provides an exception authorizing document preparation by certified nonlawyers.77 To obtain certification, an individual must satisfy the required combination of education (legal or non-legal) and legal experience, and must receive a minimum of ten hours of continuing

---

71. The exceptions in the Washington definition include:
(1) Practicing law authorized by a limited license . . . . (2) Serving as a court house facilitator pursuant to court rule. (3) Acting as a lay representative authorized by administrative agencies or tribunals. (4) Serving in a neutral capacity as a mediator, arbitrator, conciliator, or facilitator. (5) Participation in labor negotiations, arbitrations or conciliations arising under collective bargaining rights or agreements. (6) Providing assistance to another to complete a form provided by a court for protection . . . when no fee is charged to do so. (7) Acting as a legislative lobbyist. (8) Sale of legal forms in any format. (9) Activities which are preempted by Federal law. (10) Serving in a neutral capacity as a clerk or court employee providing information to the public pursuant to Supreme Court Order. (11) Such other activities that the Supreme Court has determined by published opinion do not constitute the unlicensed or unauthorized practice of law or that have been permitted under a regulatory system established by the Supreme Court.

Id.

72. Id.

73. See id.

74. See id.

75. See id.


77. Id.
education credits each year. In addition, beginning in 2005, Arizona regulations will require all nonlawyers providing legal services to pass a standardized examination.

Although Arizona’s certification process allows for greater nonlawyer activity, it does not accommodate interactive technologies. The guidelines for certification in Arizona do not explain how a nonlawyer or business may bring an interactive service into compliance. The rules enable entities to obtain certification by designating a principal who maintains valid certification as a document preparer. While this may allow nonlawyers employed by services like LegalZoom™ to obtain certification, the “document preparer” for products like Quicken® is a software program. The guidelines do not address whether these software programs may obtain certification by designating a principal. Even assuming such interactive technologies are certifiable, it is still unclear whether these services would fall outside the scope of permissible activity. Arizona’s regulation allows document preparers to provide general legal information, but does not allow them to provide specific advice, opinions, or recommendations to consumers. Because interactive technologies tailor their services to the needs of particular consumers, these programs may not comply with the Arizona regulation.

Texas is the only state with a practice of law definition that allows expansive consumer use of interactive technologies. The statute provides a broad exception to the practice of law, permitting nonlawyers to display and sell computer software and similar products if the products clearly and conspicuously state that they are not a substitute for the advice of an attorney. This exception enhances consumer access to

79. See id. at 7.
80. See id. at 6.
83. See LegalZoom™, supra note 45.
84. Tex. Gov’t Code Ann. § 81.101(c) (Vernon Supp. 2002). The Texas legislature amended the state’s original, more restrictive, definition of the practice of law in response to a federal district court decision that held that nonlawyers providing interactive legal document preparation software were engaged in the unauthorized practice of law. See Unauthorized Practice Comm. v. Parsons Tech., Inc., No. Civ. A. 3:97-CV-2859-H, 1999 WL 47235, at *1 (N.D. Tex.), vacated, 179 F.3d 956 (5th Cir. 1999); see also infra note 138 and accompanying text.
Consumer Interests & the Unauthorized Practice of Law

legal services by protecting a wide range of nonlawyer activity, presumably including products such as Quicken® WillMaker and LegalZoom™.86

Critics have raised concerns, however, regarding the statute’s sweeping assertion that software and similar products do not constitute the practice of law.87 Some note that with the advancement of technology, new software and online programs could expand the scope of the exception beyond its intended reach.88 In addition, the statute does not address reliability concerns, such as whether a nonlawyer is qualified to provide a specific service.89 Consequently, consumers must discern whether a particular product is dependable based on limited information.

2. The ABA Approach to Defining the Practice of Law Does Not Accommodate Emerging Technologies

In September 2002, a task force for the ABA, charged with the responsibility of drafting a model definition of the practice of law, released a proposed definition to the public.90 The proposed definition broadly defined the practice of law, allowing nonlawyers to provide legal services under four narrow exceptions.91 A variety of organizations instantly criticized the definition as being overly restrictive and hampering access to affordable legal services.92 Rather than revising the proposed definition, the ABA determined that a model definition was simply not a viable solution and instead issued a general report.93

In its report, the ABA recommends that all states adopt a definition of the practice of law.94 The report does not provide an example of an

86. See id.
88. See id.
89. See TEX. GOV’T CODE ANN. § 81.101.
91. The four exceptions permitted were: (1) the practice of law as authorized by a limited license; (2) pro se representation; (3) serving as a mediator, arbitrator, conciliator or facilitator; and (4) providing services under the supervision of a lawyer in compliance with the Rules of Professional Conduct. See id.
92. See Tebo, supra note 57.
93. See generally WHITSON, supra note 60.
94. See id. at 13.
effective practice of law definition, but rather urges states to balance the potential costs and benefits to consumers by considering four factors. First, states should strive towards establishing minimum qualifications that nonlawyers must satisfy in order to practice law. Second, states should ensure the competence of nonlawyers by requiring education, experience, training, certification or licensing, or supervision by a lawyer. Third, states should determine what level of professional accountability should be required of a nonlawyer. Finally, states should assess the costs and benefits associated with nonlawyer activities to determine the appropriate scope of their practice of law definitions.

Although the ABA’s suggested approach encourages states to consider important consumer protection and legal access concerns, it still advocates for a comprehensive practice of law definition. Because current practice of law definitions lack the flexibility necessary to protect consumers while preserving consumer access to beneficial nonlawyer services and emerging technologies, it is not clear that states will be able to create a definition that adequately addresses consumer interests.

C. Case Interpretation of the Unauthorized Practice of Law Does Not Provide an Exception for Interactive Technologies

States without detailed practice of law definitions place greater reliance on court interpretation of the unauthorized practice of law. The bright-line rules established by these courts have the potential to exclude consumer access to interactive technologies. Two common approaches have emerged in this area. Under the majority approach (general-specific test), nonlawyers may give legal information and advice, provided the advice is general and not tailored to an individual consumer. Nonlawyers violate the general-specific test when they

95. See id. at 5.
96. See id. at 6.
97. See id. at 7.
98. See id. at 8.
99. See id. at 9.
100. See infra notes 170–71 and accompanying text.
Consumer Interests & the Unauthorized Practice of Law

provide specific legal advice unless the activity falls within a state exception. A court following this test may decide that unauthorized practice statutes prohibit interactive technologies because products using such technologies provide specific legal advice. A minority of court decisions take a more restrictivist approach, finding that even nonlawyers providing general legal information and advice may be unlawfully practicing law in some instances. These decisions restrict consumer access to most interactive technologies. In contrast to these two approaches, Washington courts have adopted a unique approach that attempts to balance consumer protection and legal access concerns. This approach, however, does not provide set guidelines for determining when interactive technologies constitute the unauthorized practice of law.

1. The Majority General-Specific Approach Does Not Provide an Exception for Consumer Use of Interactive Technologies

The general-specific test prohibits nonlawyers from providing specific legal advice to consumers, but permits the distribution of general advice. In New York County Lawyers’ Association v. Dacey, the New York Court of Appeals overturned the lower court decision and adopted the general-specific approach proposed by the dissent. The issue in Dacey was whether distribution of a book that provided legal forms, general legal advice, and instructions to assist users in the completion of the forms constituted the unauthorized practice of law. An intermediate New York state appellate court held that under the existing New York statute, the legal information and advice provided in the book constituted the unauthorized practice of law. The dissent insisted that providing general legal advice is not the practice of law, arguing that only “personal advice on a specific problem peculiar to a designated or readily identifiable person” could constitute the practice of

103. See id.
104. See infra Part II.C.2.
105. See infra Part II.C.3.
106. Dacey, 283 N.Y.S.2d at 998 (Stevens, J., dissenting).
108. See Dacey, 234 N.E.2d at 459 (reversing the Appellate Division and dismissing the petition on dissenting opinion).
110. Id. at 995.
law. In addition, the dissent noted that the book had been sold for more than a year without any evidence to suggest consumers had been harmed or inconvenienced by the general advice.

Subsequent decisions by state courts both within and outside of New York have endorsed the general-specific test. These courts have recognized that nonlawyers providing general legal advice do not create a relationship of trust and confidence with consumers. Consequently, general legal advice does not raise the same consumer protection concerns as the distribution of specific legal advice. For example, in *Oregon State Bar v. Gilchrist*, the Oregon State Supreme Court held that it was permissible under the state’s broad practice of law statute to publish or sell divorce kits, provided the vendor did not have personal contact with their customers in the nature of consultation, explanation, recommendation, or advice. Citing *Dacey*, the court concluded that nonlawyers could render general advice on common problems.

Application of the general-specific test to interactive technologies raises new questions of interpretation. In particular, courts must determine whether interactive programs provide general or specific advice. The Oregon State Bar considered this question in a formal opinion. Referencing *Gilchrist*, the opinion suggested that a bright line exists between those who provide consumers with information in text or database form and those who personally exercise judgment in providing specific legal advice to particular individuals. According to the formal opinion, if a nonlawyer generates legal advice from a database using “decision-tree” software, it does not constitute the practice of law. However, if nonlawyers provide the same advice during an online

---

111. *Id*. at 998 (Stevens, J., dissenting).
112. *See id*. at 998–99 (Stevens, J., dissenting).
114. *See, e.g.*, *Cramer*, 249 N.W.2d at 9; *Divorce Associated*, 407 N.Y.S.2d at 144; *Gilchrist*, 538 P.2d at 917.
115. 538 P.2d 913, 917 (Or. 1975).
116. *Id*. at 913.
117. *See id*. at 917. For additional cases prohibiting nonlawyers from providing specific legal advice, see *Cramer*, 249 N.W.2d at 8–9; *Divorce Associated*, 407 N.Y.S.2d at 142.
119. *See id*. at *2.
120. *See id*.
Consumer Interests & the Unauthorized Practice of Law

session, this advice constitutes the practice of law. Thus, the Oregon State Bar’s interpretation of the general-specific test would likely prohibit online services like LegalZoom™ that use nonlawyers to prepare legal documents, and allow programs like Quicken®. Other states, however, may prohibit nonlawyers from distributing decision-tree software programs like Quicken®, finding that they too provide specific advice by tailoring legal documents to individual consumers.

2. The Minority Approach Restricts Access to Most Interactive Technologies

A minority of courts have held that even general legal information and advice provided by nonlawyers may constitute the unauthorized practice of law. Although the Texas legislature has recently adopted a less-restrictive practice of law definition, Texas cases prior to the adoption of this definition provide perhaps the most developed case law for the minority approach. In these cases, Texas courts focused on whether “legal skill or knowledge” was required to perform the service. Because all activities requiring legal skill constituted the practice of law, nonlawyers could not provide even basic legal information.

In determining whether a service required “legal skill or knowledge,” the Texas courts considered both the complexity of the area of law in which the nonlawyer was practicing and the type of legal service provided. For example, in Palmer v. Unauthorized Practice Committee, a Texas appellate court concluded that certain areas of

121. See id.
123. See TEX. GOV’T CODE ANN. § 81.101 (Vernon Supp. 2002); see also supra notes 84–85 and accompanying text.
125. See, e.g., Fadia, 830 S.W.2d at 164; Cortez, 692 S.W.2d at 50. The term “legal skill or knowledge” was adopted from Texas’ former practice of law definition. TEX. GOV’T CODE ANN. § 81.101(a) (Vernon Supp. 1998) (amended 1999).
126. See Fadia, 830 S.W.2d at 164 (citing Palmer, 438 S.W.2d at 376).
127. See, e.g., id.; see also Cortez, 692 S.W.2d at 50.
law, such as trusts, taxation, estates, and perpetuities, require such an immense amount of knowledge that they necessarily involve the practice of law requiring the skill of a licensed attorney. The nonlawyer’s lack of education and legal training compared to the complexity of the legal issues involved also concerned the court. Another Texas court focused on the activity provided, rather than the area of law at issue. In Unauthorized Practice Committee v. Cortez, the court concluded that assisting a client in the filing of certain immigration forms required legal skill because the forms furnished immigration authorities with the alien’s address, increasing the likelihood of deportation. Thus, legal skill was required to determine whether the forms should be filed.

The Texas decisions also evaluated the propensity of nonlawyer legal services to mislead consumers. In both Palmer and Fadia v. Unauthorized Practice Committee, the courts noted that advertisements and manuals implying that all testamentary dispositions could be standardized may create a false sense of security among consumers. Similarly, the Cortez court expressed concern that the nonlawyer was misleading customers by informing them that she could not help with their immigration cases without suggesting other avenues that may be available to them.

Texas’ experience with the unauthorized practice of law demonstrates the important role that specific restrictions within an unauthorized practice of law statute play in courts’ determinations of the scope of legal services nonlawyers can provide. The legal skill or knowledge test that emerged in Texas stemmed from a restrictive practice of law definition that classified legal instructions and form selection as the practice of law. Consequently, the Texas cases described above demonstrate how the language of a state’s unauthorized practice of law

129. See id. at 374.
130. See id. at 375.
131. 692 S.W.2d 47 (Tex. 1985).
132. Id. at 47, 50.
133. See, e.g., Fadia v. Unauthorized Practice of Law Comm., 830 S.W.2d 162, 165 (Tex. App. 1992); see also Cortez, 692 S.W.2d at 50.
135. See Fadia, 830 S.W.2d at 162, 165; Palmer, 438 S.W.2d at 374, 376.
136. See Cortez, 692 S.W.2d at 50.
137. See TEX. GOV’T CODE ANN. § 81.101(a) (Vernon Supp. 1998) (amended 1999); see also supra notes 124–26 and accompanying text.
Consumer Interests & the Unauthorized Practice of Law

statute or practice of law definition can limit a court’s flexibility to rule on individual cases.

Texas’ minority position prevailed until 1999, when Texas’ legislature amended the practice of law definition in response to Unauthorized Practice Committee of Law v. Parsons Technology, Inc.\textsuperscript{138} In Parsons, a federal district court enjoined the sale of Quicken\textsuperscript{®} Family Lawyer ’99 (QFL) as the unauthorized practice of law.\textsuperscript{139} QFL used decision-tree software to assist users in the selection and completion of legal forms.\textsuperscript{140} The Parsons court did not distinguish QFL because a database was generating the legal advice. Instead, the court echoed Palmer, Fadia, and Cortez\textsuperscript{141} in noting that nonlawyers cannot engage in form selection or the preparation of customized legal documents.\textsuperscript{142} The U.S. Court of Appeals for the Fifth Circuit ultimately vacated the Parsons decision\textsuperscript{143} after the Texas legislature amended the state’s unauthorized practice statute to protect access to a variety of nonlawyer services, including how-to books, kits, and interactive technologies.\textsuperscript{144}

Texas is not the only state to have applied the minority test when interpreting the unlawful practice of law. Like Texas, the Florida courts used to apply the “legal skill or knowledge” test,\textsuperscript{145} but later reevaluated their position and found persuasive reasons for adopting the general-specific test.\textsuperscript{146} Colorado continues to follow the minority approach. As recently as 1995, the Colorado State Supreme Court affirmed that the counseling and sale of “living trust” documents by a nonlawyer amounts to the unauthorized practice of law.\textsuperscript{147}

\textsuperscript{139} See id. at *6.
\textsuperscript{140} See id.
\textsuperscript{141} See supra notes 127–36 and accompanying text.
\textsuperscript{142} See Parsons, 1999 WL 47235, at *6. The court expressed concern over QFL’s packaging, which stated that its forms were valid in forty-nine states and had been updated by legal experts. The Parsons court noted that these statements created an air of reliability and could mislead consumers into depending on them. Id.
\textsuperscript{143} See Unauthorized Practice of Law Comm. v. Parsons Tech., Inc., 179 F.3d 956, 956 (5th Cir. 1999).
\textsuperscript{144} Tex. Gov’t Code Ann. § 81.101(c) (Vernon Supp. 2002).
\textsuperscript{145} See, e.g., Fla. Bar v. Stupica, 300 So. 2d 683, 687 (Fla. 1974).
\textsuperscript{146} See Fla. Bar v. Brumbaugh, 355 So. 2d 1186, 1193 (Fla. 1978).
\textsuperscript{147} See People v. Laden, 893 P.2d 771, 772 (Colo. 1995). Courts in Florida, Ohio, and Missouri have similarly held that the preparation of trust documents by nonlawyers constitutes the unauthorized practice of law. See, e.g., In re Mid-Am. Living Trust Assoc., Inc., 927 S.W.2d 855, 871 (Mo. 1996); In re Fla. Bar Advisory Opinion—Nonlawyer Preparation of Living Trusts, 613
Proponents of the minority approach insist that an expansive interpretation of the practice of law is consistent with protecting the public welfare.\textsuperscript{148} An incorrect assertion by a nonlawyer can result in substantial consumer liability, needless litigation, or unjust results.\textsuperscript{149} A seemingly simple decision, such as when to use a particular form, may involve complex legal analysis and require extensive research.\textsuperscript{150} Opponents of this approach assert that in advancing consumer protection concerns, minority states limit consumer access to nonlawyer services and threaten the availability of legal literature.\textsuperscript{151} Consequently, consumers may be restricted from using interactive technologies such as Quicken\textsuperscript{®} WillMaker or LegalZoom\textsuperscript{TM}.

3. Washington’s Approach Does Not Provide Set Guidelines for Determining when Interactive Technologies Constitute the Unauthorized Practice of Law

Washington courts have developed a unique standard for assessing the legality of nonlawyer activities. Once a court deems a nonlawyer to be practicing law, it weighs consumer protection concerns against the benefits of enhanced legal access to determine whether the activity should be permitted.\textsuperscript{152} If the court concludes that the nonlawyer service is in the public’s best interest, the activity is permitted, provided the nonlawyer adheres to the standard of care of a practicing attorney.\textsuperscript{153}

The Washington State Supreme Court discussed this approach in Jones v. Allstate Insurance Co.\textsuperscript{154} The court weighed the competing interests at stake in the case and held that although a claims adjuster was practicing law by selecting and preparing certain legal instruments, the

---

So. 2d 426, 427 (Fla. 1992); Cleveland Bar Ass’n. v. Yuriuch, 642 N.E.2d 79, 84 (Ohio Bd. Unauth. Prac. 1994).


149. See Stupica, 300 So. 2d at 687–88 ("[a]n incorrect assertion can result in perjury, libel or contempt; a warranty deed, misused for a quitclaim of limited interests, can result in substantial liability").

150. See id. at 687.


153. See id. at 305, 45 P.3d at 1075.

Consumer Interests & the Unauthorized Practice of Law

risk of harm to the public was low as compared to the benefits associated with convenient and low cost services. Consequently, the court held that claims adjusters may offer document preparation services, provided they meet the standard of care of a practicing attorney. Washington courts have also made similar exceptions for mortgage lenders and real estate brokers.

Washington’s approach is unique in that it seeks to balance consumer protection and legal access concerns. In balancing these concerns, however, Washington courts have not established a set of factors to guide their analysis. For example, in In re Disciplinary Proceedings Against Droker, the Washington State Supreme Court would not permit a nonlawyer to draft escrow instructions, select forms, draft clauses modifying form legal documents, or explain to buyers and sellers the meaning and effect of the documents they drafted, finding that these activities created too great a risk of harm to consumers. This decision and others contrast with the court’s decision in Jones to permit claims adjusters to select and complete legal documents and to advise third parties to sign such documents. Thus, it is unclear whether the courts would enjoin interactive technologies, which involve the selection and completion of legal documents and the drafting of instructions and explanations.

In sum, the development of interactive technologies by nonlawyers has forced the legal profession to seek better methods of regulating the unauthorized practice of law. Some states have responded by expanding their practice of law definitions and establishing guidelines through court interpretation. Fixed definitions and bright-line rules, however, jeopardize the flexibility necessary to keep pace with emerging technologies. At the same time, a failure to identify the factors used for

---

155. See id. at 304–05, 45 P.3d at 1075–76.
156. See id. at 305, 45 P.3d at 1075–76.
159. See id. at 719, 370 P.2d at 248.
160. See, e.g., Wash. State Bar Ass’n v. Great W. Union Fed. Sav. & Loan Ass’n, 91 Wash. 2d 48, 586 P.2d 870 (1978) (holding that the selection and completion of loan documents constituted the unauthorized practice of law).
161. See Jones, 146 Wash. 2d at 305, 45 P.3d at 1075.
interpreting the unauthorized practice of law may result in a loss of predictability and accountability in the application of these regulations.

III. A FOUR FACTOR TEST FOR COURTS TO CONSIDER WOULD PRESERVE ACCESS TO EMERGING TECHNOLOGIES WHILE PROTECTING CONSUMERS FROM UNDESIRABLE NONLAWYER ACTIVITY

The methods that states have developed for interpreting state unauthorized practice of law regulations have failed to reconcile competing legal access and consumer protection concerns. Whether a nonlawyer activity constitutes the practice of law is heavily dependent on the factual situation at hand and should not be resolved based on inflexible definitions and bright-line rules. Applying a test with specific factors to each case would better enable courts to weigh the costs and benefits of allowing a particular nonlawyer service and to adapt more effectively to innovative legal products. Under this test, courts would first determine whether the nonlawyer directed legal advice to the specific problem of an identifiable person. If so, then the court would independently weigh four factors: the reliability of the service, the potential for a conflict of interest, the service’s potential to mislead consumers, and whether the nonlawyer has addressed confidentiality and privacy concerns. The test also would ensure greater consistency and predictability among the courts. However, the test cannot provide the required flexibility if courts interpret it within the framework of a restrictive unauthorized practice of law regulation that prohibits most nonlawyer activities. Therefore, only states with less restrictive unauthorized practice of law regulations and practice of law definitions could apply this test.

162. See supra Part II.B–C.
163. See supra note 57 and accompanying text.
164. See infra notes 232–34 and accompanying text.
165. See infra notes 232–34 and accompanying text.
166. See infra Part III.B.1.
168. See infra Part III.C.
A. The Need for a Comprehensive Test

States rely on inflexible state practice of law definitions and case precedent in determining whether a nonlawyer legal service violates unauthorized practice of law statutes and court rules. An inflexible practice of law definition restricts the ability of courts and committees to assess all of the factors relevant in determining whether a particular service is beneficial to consumers. Consequently, when applied, these definitions either overly restrict nonlawyer activity or fail to provide appropriate consumer protection.

1. Inflexible Practice of Law Definitions Do Not Adequately Protect Consumer Interests

Current practice of law definitions do not successfully balance consumer protection and legal access concerns. At the expense of consumer access to legal services, most states have instituted restrictive practice of law definitions intended to protect consumers from unreliable nonlawyer services and to protect the legal profession from increased competition. Texas’ practice of law definition is at the other extreme, broadly protecting consumer access to all interactive technologies to the detriment of consumer protection.

State attempts to build exceptions into their practice of law definitions have not proven effective. Washington provides a catchall exception affording the state supreme court broad discretion to determine when the state’s unauthorized practice of law regulation should exempt a particular activity. The Washington State Supreme Court, however, has been unable to articulate a consistent standard for interpreting these nonlawyer activities. Arizona allows nonlawyers to offer such services, provided certification is obtained. But, because the certification guidelines do not allow document preparers to provide specific legal advice, Arizona’s approach may impede access to

169. See supra Part II.B–C.
170. See Christensen, supra note 6, at 161.
171. See supra notes 84–89 and accompanying text.
173. See supra notes 158–61 and accompanying text.
document services such as Quicken® WillMaker. Only Texas has a practice of law definition that protects access to a broad range of document preparation services. The Texas definition, however, protects consumer access at the expense of consumer protection. Under Texas’ statute, nonlawyers can provide legal services despite evidence that the service is unreliable.

The ABA approach encourages states to create a comprehensive definition of the practice of law, taking into consideration consumer protection and legal access considerations. No state practice of law definition, however, has achieved a constructive balance between these competing consumer interests. The more restrictive and over-particularized a practice of law definition is, the less capable it is of being adapted to new practices and legal services. The ABA itself was not able to draft a model definition that adequately balances these concerns. Consequently, it is unlikely that the ABA approach will prove successful.

2. The Insufficiency of Court Interpretation

The approaches adopted by courts for interpreting state practice of law definitions have not fared much better than restrictive practice of law definitions in protecting consumer interests. The general-specific test, favored by a majority of courts, prohibits nonlawyers from providing specific legal advice to consumers. This approach fails to recognize that it may be appropriate for a nonlawyer to provide specific advice in some circumstances. The test does not consider such factors as whether the nonlawyer is particularly well qualified to provide legal advice or whether the nonlawyer is practicing in a standardized area of law. Under the majority approach, an experienced paralegal cannot provide specific legal advice regarding an uncontested divorce, despite evidence that the paralegal is providing a valuable legal service to society.

175. See supra notes 81–83 and accompanying text.
176. See supra notes 84–86 and accompanying text.
177. See supra note 89 and accompanying text.
178. See WHITSON, supra note 60, at 13.
179. See supra notes 172–73 and accompanying text.
180. See supra notes 90–93 and accompanying text.
181. See supra note 106 and accompanying text.
182. See supra note 106 and accompanying text.
Consumer Interests & the Unauthorized Practice of Law

In addition, the general-specific test creates a bright-line rule that is inflexible and ambiguous when applied to interactive software and other technologies. Interactive programs, such as LegalZoom™, enable consumers to provide personal information to nonlawyers who use the information to customize legal forms. Because online services like LegalZoom™ provide legal advice to a “readily identifiable person” based on that person’s “specific problem,” courts can easily conclude, as the Oregon State Bar has, that these interactive services provide specific legal advice and are prohibited under state unauthorized practice statutes and court rules.

Courts applying the minority test are even more restrictive, often classifying general legal instructions as the practice of law. This approach protects consumers at the expense of consumer access to legal services by making it more difficult for consumers to obtain the legal information from nonlawyers they need to make sound choices. Because general legal advice constitutes the practice of law, nonlawyers cannot provide document preparation services such as LegalZoom™ or Quicken® WillMaker. In addition, the approach threatens the availability of literature describing relevant procedural and substantive law.

The general-specific test and the minority approach do not balance consumer protection with access to legal services. Interpretation of the general-specific test’s bright-line rule will become increasingly difficult with the development of interactive software and online technology. The minority approach has been even less successful in providing access to legal services, leading several courts to overturn decisions relying on this approach. The inadequacies of these approaches demonstrate the need for a test with flexible factors.

183. See supra notes 118–20 and accompanying text.
184. See LegalZoom™, supra note 45.
187. See supra notes 122–26 and accompanying text.
188. See supra note 148 and accompanying text.
191. See supra notes 118–19 and accompanying text.
192. See supra notes 145–46 and accompanying text.
B. A Proposed Test

Because current approaches do not adequately balance consumer protection and access to justice concerns, courts should adopt a test with four factors to determine when nonlawyer services violate unauthorized practice of law regulations. Unlike the ABA’s balancing approach, the proposed test does not attempt to define the practice of law, preferring instead courts’ application of the factors on a case-by-case basis. The test provides courts with an effective means of clarifying and interpreting state unauthorized practice of law regulations as applied to nonlawyers. States that have adopted practice of law definitions should keep these definitions as general and non-restrictive as possible. Courts would then use the proposed test to determine whether specific nonlawyer activities constitute the unauthorized practice of law. For states without practice of law definitions, the test would provide a consistent standard for classifying nonlawyer activities.

The factors selected for the test represent the overriding concerns expressed by courts in determining whether to enjoin a nonlawyer activity as a violation of an unauthorized practice of law regulation. Specifically defining these factors will help to preserve consistency and predictability among the courts. At the same time, this test still enables courts to maintain sufficient discretion to adapt to changing approaches in the law by determining which factors deserve particular emphasis.

1. Is the Legal Advice Directed to the Specific Problem of an Identifiable Person?

Initially, courts should assess whether a particular nonlawyer activity provides general or specific advice. A majority of courts have identified this threshold question as a relevant consideration. These courts have consistently held that general legal advice does not constitute the practice of law because nonlawyers providing general legal advice do not create a relationship of trust and confidence with consumers. When nonlawyers provide specific legal advice, however, the potential

194. See supra Part II.C.1.
195. See supra Part II.C.1.
Consumer Interests & the Unauthorized Practice of Law

for harm increases because consumers place greater reliance on advice specific to their case.  

While the majority approach automatically classifies services providing specific legal advice as the unauthorized practice of law, the proposed test applies additional factors before making such a determination. This approach enables courts to weigh consumer protection and legal access concerns prior to determining whether to prohibit a nonlawyer service. The proposed test protects access to beneficial nonlawyer services that involve specific legal advice, such as property transactions by real estate agents, an exception approved by several states. Changes in technology particularly warrant such an approach because emerging interactive technologies challenge the general-specific distinction.

An example of specific advice involves legal consultation or recommendation between a nonlawyer and a consumer. Books, kits, and seminars lack such a relationship and rarely involve the transmission of specific advice. The advent of online document preparation services like LegalZoom™, however, enables nonlawyers to tailor legal documents to individual consumers without assuming a personal relationship. One state has already prohibited these online services. Rather than excluding these services as the unauthorized practice of law, the proposed test would assess the activity under four factors.

2. **Is the Nonlawyer Service Reliable?**

Once courts identify a nonlawyer service as providing specific legal advice, they must then assess the service’s reliability. Several courts have identified reliability as a relevant consideration because of the devastating effects poor legal advice can have on consumers. As a New York court noted in *People v. Divorce Associated and Publishing*

---

197. See id. (noting that general legal advice raises fewer consumer protection concerns because consumers do not rely on the selection and judgment of nonlawyers providing general advice).

198. See supra note 106 and accompanying text.

199. See AM. BAR ASS’N STANDING COMM. ON CLIENT PROT., supra note 31, at 2.

200. See supra notes 106–17 and accompanying text.

201. See supra notes 106–17 and accompanying text.

202. See supra notes 118–21 and accompanying text.


204. See Tebo, supra note 4, at 41–42.
mistaken advice can jeopardize a client’s case and can lead to additional expense and litigation. Conversely, a demonstrably reliable service does not raise the same concerns. The dissent in \textit{Dacey} recognized as much, noting that the book at issue had been published and sold for more than a year without a showing that it had exploited the public or led its members astray. Thus, the proposed test should account for reliability in assessing whether a product violates state unauthorized practice of law regulations.

In assessing the reliability of a particular product or service, courts should focus on a number of factors. First, they should consider the education and experience of the nonlawyer. In \textit{Palmer}, the court noted that the nonlawyer had not completed high school and was untrained in the law. A showing of significant legal experience, however, may overcome a lack of education. In Arizona, a high school diploma is sufficient for certification as a legal document preparer when combined with a minimum of two years of law-related experience. Education and experience are appropriate considerations because they demonstrate to the court whether a nonlawyer has the necessary skills to provide a legal service independently.

Second, a court should evaluate the complexity of the nonlawyer service. For example, if the service involves a complicated area of law, it is more likely that someone without expertise in that area will be unreliable. \textit{Palmer} noted that the areas of trust, tax, and estate law are highly complicated, requiring additional expertise. In contrast, as the court noted in \textit{Florida Bar v. Brumbaugh}, certain areas of law such as the uncontested dissolution of marriage are less complex and have more easily standardized services, raising fewer reliability concerns.

\begin{footnotesize}
206. See \textit{id.} at 144–45.
210. See \textit{id.}
211. See \textit{Palmer}, 438 S.W.2d at 376.
212. 355 So. 2d 1186 (Fla. 1978).
213. See \textit{id.} at 1193.
\end{footnotesize}
Consumer Interests & the Unauthorized Practice of Law

Third, a court should establish whether a licensed attorney has reviewed the legal product or service for legal accuracy. Particularly when the nonlawyer lacks significant education or expertise, supervision by an attorney suggests greater reliability of the service. Finally, as expressed by the Divorce Associated court, an indication that the product is erroneous and inadequate would demonstrate the product’s unreliability. If, after evaluation of these factors, a court deems the product or service to be unreliable and to pose a threat to consumers, such a finding would be sufficient for the court to enjoin the activity as the unauthorized practice of law.

3. Is There Potential for a Conflict of Interest?

Courts should next consider whether a particular nonlawyer legal service presents a conflict of interest. A conflict of interest may arise if a nonlawyer provides document preparation services or specific legal advice to directly adverse parties. Nonlawyers may also have personal interest conflicts. For example, banks and trust marketing corporations that stand to benefit from the sale of legal instruments face conflicting interests when providing clients with specific legal advice regarding estate-planning services.

Courts should assess nonlawyer conflicts of interest in accordance with the applicable state rules of professional conduct for attorneys. Because nonlawyers create a relationship of trust and confidence when providing specific legal advice to consumers, states should hold them to the same conflict of interest standard as attorneys. Holding nonlawyers performing legal services to this same standard would protect consumers against the most serious conflicts of interest. A court

214. See Fadia v. Unauthorized Practice of Law Comm., 830 S.W.2d 162, 165 (Tex. App. 1992) (noting that the defendant had not had his will manual reviewed by a licensed attorney).
216. See supra note 21 and accompanying text.
217. See In re Mid-Am. Living Trust Assoc., Inc., 927 S.W.2d 855, 860 (Mo. 1996) (noting that “a conflict of interest exists between those who benefit from the sale of a particular legal instrument and the client for whom that legal instrument may not be appropriate”).
218. See MODEL RULES OF PROF’L CONDUCT R. 1.7 (2002).
may enjoin nonlawyer activities that fail to follow these guidelines for violating state unauthorized practice of law regulations.

4. Does the Service Have the Potential To Mislead Consumers?

A court should also assess the likelihood that a product’s advertising, disclaimers, or advice will mislead consumers. Erroneous specific legal advice can seriously harm consumers. The Texas decisions recognized the importance that nonlawyers accurately portray the inherent risks of their services and not overstate their abilities. For example, the Palmer and Fadia courts found that a combination of the advertisements and the product itself misled consumers by creating the false impression that all testamentary dispositions may be standardized. A product may also create false expectations as to its reliability. The Parsons court found that a product’s packaging increased the likelihood that consumers would be misled into relying on the product by advertising that its forms were valid in forty-nine states and had been reviewed by legal experts. The inadequacy of the product’s disclaimer, which only actively appeared the first time a consumer used the program, compounded the potential for false impressions.

If nonlawyers are seriously misleading consumers by misrepresenting their services or the law, a court should find that consumer protection concerns outweigh the product’s potential benefits and enjoin the activity. A comprehensive disclaimer may remedy minor misconceptions, but it should not protect nonlawyers that critically mislead consumers.

5. Has the Nonlawyer Addressed Confidentiality and Privacy Concerns?

The final factor that courts should consider is whether the product or service addresses confidentiality concerns. States have recognized limited confidentiality protections between nonlawyers and clients in

221. See Tebo, supra note 4, at 41–42.
222. See supra notes 133–36 and accompanying text.
225. See id.
Consumer Interests & the Unauthorized Practice of Law

some instances. For example, tax advice, communicated between a
taxpayer and a federally authorized accountant-tax practitioner, is
subject to general protections of confidentiality.\(^{226}\) In general, however,
the protection that clients receive under the attorney-client privilege is
not available when clients provide personal legal information to
nonlawyers.\(^{227}\) As a result, courts may compel nonlawyers to testify
about communications involving the client.\(^{228}\) Courts should require that
nonlawyers disclose this risk to clients at the outset of the relationship.
Nonlawyers could satisfy this requirement by way of a written
disclaimer, provided the disclaimer is unambiguous and is shown to all
clients prior to use of the product or service. If consumers solicit a
nonlawyer’s services after reading the disclaimer, then they are
presumed to have impliedly consented to the lack of privilege, much like
consumers can waive their right to confidentiality with their lawyers.\(^{229}\)

Nonlawyers should also address basic privacy concerns by providing
consumers with comprehensive privacy policies and ensuring that
interactive web-based products have proper data encryption and security
devices. The ABA Model Rules of Professional Conduct require that
attorneys take “reasonable precautions to prevent . . . information from
coming into the hands of unintended recipients,” while noting that this
duty “does not require that the lawyer use special security measures if
the method of communication affords a reasonable expectation of
privacy.”\(^{230}\) Because a breach of confidentiality may cause
embarrassment or prove legally damaging to the consumer,\(^{231}\)
nonlawyers providing legal services should be held to a comparable
standard as that applied to lawyers in a particular state. Thus, in states
that have adopted the ABA Model Rules, nonlawyers should be required
to take reasonable precautions to prevent the careless disclosure of
sensitive information. If a product or service fails to reasonably protect
consumers’ privacy and confidentiality, or consumers are not informed

\(^{226}\) See HAZARD, supra note 220, at 226.

\(^{227}\) See supra note 26 and accompanying text.

\(^{228}\) See, e.g., Hunt v. Maricopa County Employees Merit Sys. Comm., 619 P.2d 1036, 1041
(Ariz. 1980).

\(^{229}\) See HAZARD, supra note 220, at 271 (noting that the Model Rules of Professional Conduct
allow for the disclosure of any client confidence if the client consents after consultation with the
lawyer).

\(^{230}\) MODEL RULES OF PROF’L CONDUCT R. 1.6 cmt. (2002).

\(^{231}\) See id. (noting that clients are encouraged to provide embarrassing or legally damaging
information to their attorneys).
of the lack of an attorney-client privilege, the court should enjoin the activity for violating unauthorized practice of law regulations.

C. Application of the Test

The proposed test would provide courts with a more flexible method of interpreting unauthorized practice statutes and court rules for two reasons. First, adoption of the test would enable courts and committees to consider a variety of factors in determining whether a particular nonlawyer activity violates an unauthorized practice of law regulation. These factors evaluate the benefits that consumers receive from the service as well as the risks associated with its use, allowing courts to base their decisions on the consumers’ interests rather than on the classification of the type of service being provided. Second, the flexibility of the test would permit courts to apply the standard to new technologies without having to alter the structure of the test—a major benefit over all-encompassing definitions, which must continuously change to incorporate new services into their system of classification.

The proposed test does not compromise the predictability of unauthorized practice regulations. By first asking whether a practice complies with the general-specific test’s bright-line rule, the test incorporates a standard already in use by the majority of state courts. Nonlawyers providing general legal advice are under the same standards as they were with the general-specific test. The proposed test simply enables nonlawyers to provide certain specific legal advice if they are able to satisfy the four factors.

Although these additional considerations allow for greater discretion than a bright-line rule, they still provide nonlawyers with a predictable framework for determining whether a particular activity constitutes the unauthorized practice of law. Lawyers interpret and abide by these factors in practice. There is no reason to suspect that nonlawyers

234. See supra note 57 and accompanying text.
235. See supra notes 194–96 and accompanying text.
236. See supra notes 198–99 and accompanying text.
237. The Model Rules of Professional Conduct address concerns regarding reliability, conflicts of interest, misrepresentation, and confidentiality. See, e.g., MODEL RULES OF PROF’L CONDUCT R. 1.1, 1.6, 1.7, 4.1, 7.2 (2002). Lawyers in a majority of states are regulated by some version of these rules. See HAZARD, supra note 220, at 115.
could not do the same because they would be aware of the standards they must meet to practice law. And while the proposed test affords the courts with greater discretion in determining which activities constitute the unauthorized practice of law, it also limits this discretion by requiring courts to support their conclusions based on the four factors.

Washington and Texas serve as two useful examples of how to incorporate the proposed test into a less restrictive practice of law definition. Washington established its practice of law definition by court rule, providing a catchall exception for activities that the state’s supreme court determines do not constitute the unauthorized practice of law. Washington State Supreme Court currently evaluates nonlawyer activities that fall within this exception by weighing consumer protection and legal access concerns. However, the court has been inconsistent in balancing these competing concerns and has not established guidelines that nonlawyers can follow. Adopting the proposed test would provide the court with a set of factors to adhere to in determining when to permit a nonlawyer to practice law, promoting greater consistency and predictability among Washington State Supreme Court opinions.

Texas’ practice of law definition is a statutory provision that allows for the display and sale of computer software and similar products if those products clearly and conspicuously state that they are not a substitute for the advice of an attorney. Although this broad exception to the practice of law preserves legal access to a variety of nonlawyer services, the exception does little to protect consumers from harmful nonlawyer products. Texas should adopt the proposed test as a subset of the state’s practice of law statute, so that nonlawyers providing computer software and similar products remain subject to assessments of reliability, misrepresentation, conflicts of interest, and confidentiality.

IV. CONCLUSION

State courts and bar committees have struggled to create an all-encompassing rule that describes when a nonlawyer activity constitutes the practice of law. The result has been a myriad of inconsistent practice of law definitions and unauthorized practice decisions that fail to strike

238. See WASH. CT. GEN. R. 24.
240. See supra notes 158–61 and accompanying text.
an appropriate balance between consumer protection and legal access concerns. Rather than attempting to define the practice of law, the proposed test provides a set of four factors that courts and committees should apply within the framework of less-restrictive state unauthorized practice regulations or practice of law definitions to determine whether nonlawyers are engaged in the unauthorized practice of law. These factors are central to the unauthorized practice debate. Incorporating them into the proposed test provides courts and committees with a crucial tool for determining when nonlawyer services serve consumers’ interests.
SEX AND THE WORKPLACE: “CONSENTING” ADOLESCENTS AND A CONFLICT OF LAWS

Jennifer Ann Drobac∗

Abstract: According to the Equal Employment Opportunity Commission, sexual harassment of adolescents at work may constitute a serious, but to date largely undocumented, problem. Courts respond inconsistently to adolescent “consent” in sexual harassment employment cases. This Article reviews state criminal statutory rape law, federal civil law, and tort law to reveal the conflicting legal treatment of adolescent capacity to consent to sex. It highlights conflicts not only between the criminal and civil systems, but also between sister states’ laws and laws within states. For example, this Article finds that despite criminal sexual abuse laws, courts permitted employers to use adolescent “consent” as a defense to sexual harassment in approximately fifty percent of the surveyed common law tort cases across the nation. After exploring the public policy goals for these various laws, this Article concludes that these goals do not justify the blatant conflicts between tort and criminal laws. This Article recommends both administrative and statutory reform to protect minors from the predation of adult supervisors and employers. Particularly, it recommends a strict liability standard in an approach that makes an adolescent’s consent to sex with an adult at work voidable by the minor.

When Sara was fifteen, the forty-year-old manager of the movie theater where she worked befriended her and gained her confidence. His attention became increasingly intimate and physical.1 At first, she rebuffed his physical advances. After several months, he told her that he had a brain tumor and was not sure how long he had to live. He told her he loved her. When she was sixteen, she agreed to have sexual intercourse with him. He promoted her to projectionist so that they could engage in sex more easily and frequently in the secluded projection room. She was soon pregnant. His adult girlfriend took her to have an abortion. Her manager was already in jail, serving time on a larceny conviction. Sara’s parents knew nothing about the affair. Sara wrote to

∗ Associate Professor of Law, Indiana University School of Law-Indianapolis. J.S.D., Stanford Law School, 2000. My sincerest gratitude to my readers, Martin Drobac, Esq., Professors Cynthia Baker, Dan Cole, Robin Craig, Kenneth Crews, Nicholas Georgakopoulos, Andrew Klein, Deborah Malamud, Antony Page, Florence Roisman, Elaine Sutherland, and R. George Wright, who reviewed and commented upon earlier drafts. Special thanks to my research assistants, Miriam Murphy, Associate Director of our library, Margaret Marr, Esq., and Sandie McCarthy-Brown for their diligent efforts. I also thank Articles Editor Patrick McKenna for his comments and work. Finally, I am grateful for summer research stipends from Indiana University School of Law-Indianapolis that supported this work.

1. Complaint at 3–9, Sara Does v. Culver Theaters (Santa Cruz Super. Ct. 1999) (No. CV139513) [hereinafter Sara Doe Complaint].
him, believing he was wrongly convicted. When Sara’s parents finally discovered the cause of her plummeting grades and disturbing behavior, they notified police, who told Sara the man had no tumor and was a registered sex offender. Sara cooperated with the district attorney, who prosecuted him for statutory rape.

Does Sara have a justiciable claim for sexual harassment against the theater owner? This Article explores that question by examining her case within different legal regimes in numerous states.

This Article explores the laws of several states that address workplace sexual harassment of adolescents. In particular, it examines the legal capacity of teenagers to consent to sexual conduct with a co-worker or supervisor.

Following Part I's introduction of adolescent legal rights and employment, Part II compares the treatment of adolescent “consent” in criminal statutory rape prosecutions with multiple interpretations of the same “consent” in civil disputes. A comparison of state statutory rape laws, Title VII of the Civil Rights Act of 1964 (Title VII), state fair employment practice statutes (FEPS), and state personal injury laws reveals the grossly inconsistent treatment of teenage consent between the criminal and civil law systems in the United States.

Part III discusses the public policy reasons for particular approaches to adolescent “consent.” This review of the goals of criminal statutory rape laws and related civil laws highlights the overlapping functions of each system. It demonstrates that public policy goals may not explain the varied treatment of adolescent “consent.”

---

2. Professor Deborah Malamud suggested that this is the wrong question with which to start. Conversation with Deborah Malamud, at the Law & Society Conference, Pittsburgh, Pa. (June 7, 2003). She suggested a regulatory approach to this problem to prevent harassment of minors in the first place. Id. I agree that a regulatory approach could prove fruitful, if combined with litigation and statutory reform. I explore the possible regulatory approaches in Part V of this Article.

3. While the term “teenager” technically includes eighteen- and nineteen-year-olds, I focus on minors in this Article. I also use the term “adolescents” when referring to minors, even though new research indicates that adolescence continues into the early twenties. See Jennifer Ann Drobac, I Can’t to I Kant: The Transition to Maturity and the Meaning of Adolescent Consent in the Workplace 8–9 (2004) (unpublished manuscript, on file with author). Additionally, I refer to both male and female teenagers, even when I use the female pronoun, because both males and females experience sexual harassment at work.

4. I use quotations with adolescent “consent” because even explicit verbal consent by a minor may not constitute legal consent and may equate more realistically with acquiescence. See Meritor Sav. Bank v. Vinson, 477 U.S. 57, 68 (1986) (holding that acquiescence is not consent in an evaluation of the "unwelcomeness" of sexual conduct under Title VII).

Adolescent Sex & the Workplace

Part IV surveys the civil case law that addresses adolescent “consent” to sexual conduct. This Part analyzes how courts interpret the law and how their decisions support or oppose public policy goals. It concludes that legal treatment of Sara’s consent, and her likelihood of prevailing in a suit for sexual harassment, depends upon the state in which she consented and files suit, and upon which claims she alleges.

Finally, Part V offers a synthesis of current law and public policy. Concluding that public policy goals fail to justify the blatant conflict in the law, this Part explores several options for future treatment of adolescent “consent” in the workplace. This Part also investigates possible regulatory approaches, as well as litigation and statutory reform. It discusses the advantages and disadvantages of each approach and offers a final recommendation: a strict liability standard that makes adolescent “consent” to sex with an adult at work voidable by the minor. In sum, this Article suggests how the law might permit Sara to sue for sexual harassment, whether she “consented” or not.

I. ADOLESCENT LEGAL RIGHTS & EMPLOYMENT

A. Adolescent Legal Rights

The law treats adolescents differently than it does adults. The U.S. Supreme Court confirmed in Planned Parenthood of Central Missouri v. Danforth that “[c]onstitutional rights do not mature and come into being magically only when one attains the state-defined age of majority. Minors, as well as adults, are protected by the Constitution and possess constitutional rights.” However, in Thompson v. Oklahoma, the Court acknowledged:

“[T]here are differences [between children and adults] which must be accommodated in determining the rights and duties of children as compared with those of adults. Examples of this distinction abound in our law: in contracts, in torts, in criminal

6. All states but four set the age of majority at eighteen. In Alabama and Nebraska, persons reach their majority at nineteen. In Pennsylvania and Mississippi, the age is twenty-one. Heather Boonstra & Elizabeth Nash, Minors and the Right To Consent to Health Care, 3 THE GUTTMACHER REP. ON PUB. POL’Y NO. 4, at 7 (The Alan Guttmacher Inst.), at http://www.agi-usa.org/pubs/journals/gr030404.pdf.
8. Id. at 74.
law and procedure, in criminal sanctions and rehabilitation, and
in the right to vote and hold office.\textsuperscript{10}

The Thompson Court noted many of the legal limitations on the rights of
minors, including eligibility to vote,\textsuperscript{11} to serve on a jury,\textsuperscript{12} to marry\textsuperscript{13} and
drive\textsuperscript{14} without parental consent; to purchase alcohol,\textsuperscript{15} pornographic
materials,\textsuperscript{16} and cigarettes,\textsuperscript{17} and to gamble.\textsuperscript{18}

The Thompson Court’s explanation for the different legal treatment of
minors relied in part on legal precedent\textsuperscript{19} but also on then-current studies

\begin{enumerate}
\item \textsuperscript{10} Id. at 823 (quoting Goss v. Lopez, 419 U.S. 565, 590–91 (1975) (Powell, J., dissenting)).
\item \textsuperscript{11} The U.S. Constitution grants eighteen-year-olds the right to vote. U.S. CONST. amend. XXVI. The Thompson Court noted that no state had lowered its voting age below eighteen. 487 U.S. at 839.
\item \textsuperscript{12} In Thompson, the Court noted that no state had granted this right to persons under eighteen. 487 U.S. at 840.
\item \textsuperscript{13} Id. at 824. Only Mississippi permits persons younger than eighteen to marry without parental consent or judicial authorization. See Legal Info. Inst., Marriage Laws of the Fifty States, District of Columbia and Puerto Rico, at http://www.law.cornell.edu/topics/Table_Marriage.htm (last visited Apr. 20, 2004). Seven states (Delaware, Florida, Hawaii, Indiana, Kentucky, Maryland, and Oklahoma) allow minors to marry if they are pregnant or have a child. Boonstra & Nash, supra note 6, at 6–7.
\item \textsuperscript{14} The Thompson Court noted that in all states but one, the minimum requirement for a driver’s license without parental consent was at least sixteen. 487 U.S. at 842.
\item \textsuperscript{15} Id. at 823. By 1988, all states had established a legal minimum drinking age of twenty-one. ROBERT H. MOOKIN & D. KELLY WEISBERG, CHILD, FAMILY, AND STATE 1087 (4th ed. 2000). In 1984, Congress amended federal law to withhold highway construction funds from states that failed to impose a minimum drinking age of twenty-one by 1986. Id.; see 23 U.S.C. § 158(a)(2) (1994). Research confirms that a disproportionately high percentage of fatal car accidents still involve teenagers. See Mookin & Weisberg, supra, at 1099–1100; NAT’L HIGHWAY TRAFFIC SAFETY ADMIN. (NHTSA), U.S. DEP’T OF TRANSP., TRAFFIC SAFETY FACTS (1997), at http://braininjuryoklahoma.org/intro/Transportation/traffic%20safety%20facts%201997.htm#sumar y. Mookin and Weisberg explain, “As a group, teenagers have had the least amount of experience with either activity [drinking or driving] and thus are more likely to misjudge their abilities and reactions. They are affected by small amounts of alcohol to a greater degree than more experienced drinkers.” Mookin & Weisberg, supra, at 1100.
\item \textsuperscript{16} The Thompson Court stated that no state allowed a minor to purchase obscene materials. Thompson, 487 U.S. at 845; see also Ginsberg v. New York, 390 U.S. 629 (1968).
\item \textsuperscript{18} The Thompson Court explained that thirty-nine of the forty-eight states that permitted gambling prohibited participation by minors. Another three prohibit it without the consent of the parents. Thompson, 487 U.S. at 847.
\item \textsuperscript{19} See id. at 834 (citing Eddings v. Oklahoma, 455 U.S. 104, 115–16 (1982); Bellotti v. Baird, 443 U.S. 622, 635 (1979)). The Bellotti Court noted three reasons for limiting the rights of children: “the peculiar vulnerability of children; their inability to make critical decisions in an informed, mature manner; and the importance of the parental role in child rearing.” 443 U.S. at 634.
\end{enumerate}
Adolescent Sex & the Workplace

regarding adolescent psychosocial development. The Court concluded that “[i]nexperience, less education, and less intelligence make the teenager less able to evaluate the consequences of his or her conduct while at the same time he or she is much more apt to be motivated by mere emotion or peer pressure than is an adult.”

The law limits adolescents in many other respects not covered in the Thompson decision. For example, common law declared that contracts with a minor were not void, but were voidable by the minor. This law remains the majority rule. In Stanford v. Kentucky, the dissent noted that thirty-seven states restrict a minor’s access to general medical treatment. Many states prohibit minors from filing lawsuits unless represented by a parent, next friend, or guardian. Additionally, the law limits the rights of adolescents to work. The federal Fair Labor Standards Act specifies a minimum work age of fourteen for non-agricultural work and child labor standards.

20. Thompson, 487 U.S. at 835 n.43. For a more recent discussion of adolescent psychosocial development as it relates to sexual harassment law, see Drobac, supra note 3, at 14–16, 22–26.


25. Id. at 394 (Brennan, J., dissenting); see In re Stanford, 537 U.S. 968, 969–70 (2002) (Stevens, J., dissenting). Currently, parents retain the legal authority to consent to medical treatment for their children “based on the principle that young people generally lack the maturity and judgment to make fully informed decisions before they reach the age of majority.” Cynthia Dailard, New Medical Records Privacy Rule: The Interface with Teen Access to Confidential Care, 6 THE GUTTMACHER REP. ON PUB. POL’Y NO. 1, at 6 (The Alan Guttmacher Inst. 2003), at http://www.guttmacher.org/pubs/journals/gr060106.pdf.

26. See, e.g., Porter v. Triad of Arizona, 52 P.3d 799, 803 (Ariz. Ct. App. 2002) (holding that a minor may not bring an action in his own name but may sue through a representative); Am. Alternative Energy Partners II v. Windridge, Inc., 49 Cal. Rptr. 2d 686, 690–91 (Cal. Ct. App. 1996) (finding that the incapacity of minors bars them from representing their own interests in court); Newman v. Newman, 663 A.2d 980, 987 (Conn. 1995) (holding that a child may bring an action only through a next friend or guardian); Cleaver v. George Staton Co., 908 S.W.2d 468, 669 (Tex. Ct. App. 1995) (finding a lack of capacity because of the disability of minority pertaining to the right to sue in one’s own name); Jensen ex rel. Stierman v. McPherson, 655 N.W.2d 487, 491 (Wis. Ct. App. 2002) (relying on Wis. STAT. § 803.01(3)(c) that requires that an adult represent the minor); see also Klak v. Skellion, 741 N.E.2d 288, 298–90 (Ill. App. Ct. 2000) (stating that a minor has no capacity to maintain an action in his name).

Not all laws treat adolescents as “infants” or as children with a legal
disability. Some laws grant adolescents adult privileges. For example,
all fifty states allow minors to consent, without parental approval or
notification, to testing for HIV and sexually transmitted diseases (STDs).
Forty-seven states permit minors to consent to treatment for STDs. Thirty-four states permit a minor mother to place her child for adoption
without consulting her own parents. States also attribute adult
responsibilities to adolescents. For example, California automatically
tries a minor of fourteen or older as an adult for the crimes of murder,
rape, and certain other sex offenses.

This very brief survey of adolescent rights and responsibilities raises
some interesting questions about the relationship between adolescent
employment and sexual activity. We know that the law permits
adolescents to work, albeit with limitations. Additionally, they can
consent to some medical services related to sexual activity, which means
that the law acknowledges that minors engage in sex. However, almost
seventy percent of U.S. adults believe that adolescents fourteen to
sixteen years old should not have sex.

visited Apr. 20, 2004). Federal law allows employers to pay workers under twenty less than
minimum wage for the first ninety days of their employment. U.S. Dep’t of Labor, Wages, Youth &

28. See generally Boonstra & Nash, supra note 6 (providing a useful survey of a minor’s right in
all fifty states to consent to various types of health care).

29. Id. at 5.

30. Id. at 8.

31. Dr. Michael J. Bradley discusses the trend to perceive adolescents as adults in his recent book.
MICHAEL J. BRADLEY, YES, YOUR TEEN IS CRAZY! (2002). He writes:

[We’ve somehow come to view adolescents as if they were adults and not children. From the
kid’s perspective, this is nothing new. Teenagers of all generations have lobbied for adult
privileges with the swaggering assurances that they can handle “it.” The fact is that they cannot
handle “it” and they know this . . . . What’s new is that we’ve somehow signed onto this
disastrous notion that they are adults, capable of handling “it” completely solo. It’s not
working. Teens left on their own as small adults not only screw up big-time, they become
depressed and rageful in the bargain.

Id. at 16–17.

32. CAL. WELF. & INST. CODE § 707(d)(2), (c) (West 1998). For a discussion of how states treat
juvenile criminal offenders as adults, see Elizabeth S. Scott, The Legal Construction of Adolescence,

33. For a discussion of recent statistics regarding teenage sexuality and sexual harassment, see Drobac, supra note 3, at 16–25.

34. Jennifer J. Frost et al., The Alan Guttmacher Institute, Country Report for the United States:
In Teenage Sexual and Reproductive Behavior in Developed Countries, at 18 (Occasional Report
Adolescent Sex & the Workplace

harassment a problem for minors? May minors legally consent to workplace sex, or does the law penalize them, or others, if minors engage in sex with an adult on the job? Can minors sue, through a friend or guardian, for workplace sexual harassment if they do “consent” to sex? More specifically, is Sara’s case aberrational, and if not, how will the law treat her circumstances?

B. Adolescent Employment

Researchers estimate that approximately five-and-a-half million adolescents between the ages of twelve and seventeen work. Approximately one-third of working adolescents find jobs in eating and drinking establishments. Another third work in other retail jobs. Approximately twenty-five percent work in service industries, including health, education, and entertainment or recreation jobs. Working in a movie theater, Sara fit into this last category.

1. Youth Employment Generally

Particular individual characteristics correlate with youth employment. Fifty percent of employed American high school students work more than fifteen hours per week. Black, Hispanic, and poor youth find less work and are unemployed at higher rates than their white and wealthier peers. Similarly, adolescents in single-parent families experience more unemployment than those in married-couple families.

28, 2004) (reporting that “[i]n 1996, 69% of U.S. adults indicated that it was ‘always wrong’ for adolescents between the ages of 14–16 to have sex”).


37. Youth Employment Statistics, supra note 35.


39. Id. at 33. Adolescents in married-couple families experienced fifteen percent unemployment during the 1996–1998 period as compared to almost double that (twenty-nine percent) for those in
unemployment rates typically far exceed those for other groups, averaging between fifteen and thirty-five percent, depending on age, race, family type, and time of year. Most recent figures from 2002 concerning sixteen- and seventeen-year-olds show that while approximately twenty-seven percent were employed, another nineteen percent sought work. The question arises whether adolescent workers, because they experience more unemployment than adults, are therefore more susceptible than their adult co-workers to sexual abuse and coercive demands by employers and supervisors.

40. Id. at 31. The annual average unemployment rate of fifteen to seventeen-year-olds during the 1996–1998 period hovered at nineteen percent. Id. at 33. School-month unemployment averaged at seventeen percent for whites, thirty-five percent for Blacks, and thirty percent for Hispanics. Id. at 31. Male unemployment rates were slightly higher than those for females, twenty percent compared to seventeen percent. Id.


42. Professor Robin Craig notes that adolescent unemployment may not carry the same import as adult unemployment and suggests that I may not be exploring the right question here. In a May 2003 presentation, I suggested to Professor Craig and other colleagues that new research concerning adolescent neurological and psychosocial development may dramatically change what we “know” about adolescence. In a draft paper based upon that research, I summarize that “[c]ritical abilities—including impulse control, emotional regulation, planning, decision-making, and organization—may not fully mature until the third decade of life.” Drobac, supra note 3, at 9. I believe, and Professor Craig shares the view, that adolescent development makes minors more vulnerable than adults to workplace sexual harassment. I explored this notion more fully in that paper. See id. at 4–26.

Thus, Professor Craig wonders whether adolescents may suffer higher rates of sexual harassment not because of associated high unemployment rates but because of these maturational vulnerabilities and limited experiences in the workplace. She is skeptical of the importance of adolescent unemployment rates because people, thinking that adolescents should be in school and not at work, may discount those statistics. Moreover, most minors are dependents. Because they are not responsible for paying for their own necessaries, Professor Craig suggests that work is less critical to them and, by implication, these adolescents will tolerate less harassment. I think both maturational vulnerabilities and high unemployment rates may influence the prevalence of the sexual harassment of teen workers. In any case, the significance of adolescent unemployment and its relevance to sexual harassment deserves further investigation.
2. Youth Sexual Harassment in the Restaurant Industry

Equal Employment Opportunity Commission (EEOC) charge\textsuperscript{43} statistics tend to support the notion that those industry sectors that employ more teen workers receive a high percentage of EEOC discrimination charges. For example, in 2002, the EEOC received 268 sexual harassment charges from workers under eighteen.\textsuperscript{44} Of those, forty-six percent (123) came from adolescents who worked in eating and drinking establishments.\textsuperscript{45} This statistic marks more than a ten percent increase in five years over the 1998 figure, in which only thirty-five percent of the charges came from teen workers in this industry.\textsuperscript{46}

EEOC lawsuits document sexual harassment of youth in the restaurant industry. In 2002, the EEOC filed eighteen sexual harassment lawsuits against restaurants,\textsuperscript{47} including Taco Bell, Church’s Chicken, Applebee’s, Denny’s, Colonial Ice Cream, Pepe’s Mexican Restaurant, and others.\textsuperscript{48} One cannot always tell from EEOC press releases whether the alleged targets were minors at the time the incidents occurred; nevertheless, two EEOC press releases from 2003 documented the sexual harassment of female teenagers at Burger King, Church’s

\begin{footnotes}
\item[43] I use the term “charge” rather than “complaint” to distinguish EEOC administrative charges (filed to initiate EEOC investigations) from lawsuit complaints (filed to initiate a court action).
\item[44] ORIP/PPAD, EEOC, EEOC/FEPA Charge Trends for 17 and Under Sexual Harassment Receipts by Charging Party Group—National (EEOC computer-generated statistics compiled on May 27, 2003) (on file with author) [hereinafter EEOC National 17 and Under Receipts]. I thank EEOC administrators and employees who provided me with the information discussed in this section and the footnotes that follow.
\item[46] EEOC National 17 and Under Receipts, \textit{supra} note 44; EEOC Eating and Drinking 17 and Under Receipts, \textit{supra} note 45. According to EEOC trial attorney Sanya Hill Maxion, the EEOC processed 1281 sexual harassment cases filed by workers eighteen or under in 1994. Alexei Oreskovic, \textit{Targeted Teens}, \textit{The Recorder}, July 15, 2003, at 1. The Oreskovic article also states that for the year 2003, minors had filed 541 cases, ninety-nine percent of which involved sexual harassment. \textit{Id}. The EEOC statistics provided by the Commission refute that estimate: as of May 27, 2003, minors had filed only 133 sexual harassment cases in 2003. EEOC National 17 and Under Receipts, \textit{supra} note 44.
\end{footnotes}
Chicken, Jack-in-the-Box, and Taco Bell. According to EEOC District Director Lynn Y. Bruner, no other industry receives as many complaints or sees as many cases filed in the civil court system. According to EEOC District Director Lynn Y. Bruner, no other industry receives as many complaints or sees as many cases filed in the civil court system. Is there a correlation between teen employment and sexual harassment charges more generally? The answer to this question remains unclear. We do know that in recent years, workers filed more than 1200 discrimination charges annually against restaurants, the sector that employs one-third of all adolescents. Thus, we know that minors filed approximately ten percent of the eating and drinking industry sexual harassment charges. The problem is that we do not know exactly what percentage of the industry workforce teenagers comprise.

New anecdotal evidence of the sexual harassment of adolescent workers, particularly in the fast food industry, has alerted EEOC officials to what may be an alarming phenomenon and trend. For example, on January 31, 2003, the St. Louis Post Dispatch reported the case of a Burger King manager who sexually harassed six women, five of whom were high school students. The manager, in his late twenties, ...

49. This case alleged the rape of a fourteen-year-old girl by her store manager. Sanchez, supra note 47, at 8.
51. Sanchez, supra note 47, at 8.
52. Berta, supra note 48.
53. The U.S. Department of Labor reports that 2,573,000 sixteen- and seventeen-year-olds worked in 2000. U.S. Dep’t of Labor, Household Data Annual Averages, in CURRENT POPULATION SURVEY (2000), at http://www.bls.gov/opub/rylf/rylfhome.htm. It does not report the work statistics of fourteen- and fifteen-year-olds. Id. The number of sixteen- and seventeen-year-olds represented approximately two percent of the total employed workforce. Id. I could not find Department statistics that break down that figure by industry. If, however, the percentage holds for the eating and drinking industry (which it may not since that sector employs one-third of all teenagers), then two percent of the workers filed ten percent of the sexual harassment charges.
54. Telephone Interview with David Grinberg, Spokesperson, EEOC (May 22, 2003); E-mail from William Tamayo, Regional Attorney, EEOC, San Francisco to author (May 22, 2003, 7:06 p.m. CST) (on file with author) (“Our office has litigated a few cases involving teenagers subjected to sexual harassment, and we see a growing trend.”); see also Oreskovic, supra note 46, at 1 (“Among the millions of teenagers who staff the nation’s fast food and retail outlets, sexual harassment is a pervasive problem that’s long existed under the radar.”). EEOC Program Analyst Linda Li notes that the agency “has not consistently tracked the age of the harassment victims.” E-mail from Linda Li, Program Analyst, EEOC, to author (May 22, 2003, 5:18 p.m. CST) (on file with author). Growing awareness of this phenomenon may cause more detailed tracking of the sexual harassment of adolescent workers.
Adolescent Sex & the Workplace

allegedly fondled the workers, made vulgar comments, and demanded sex. Initially, the women did not know how to make a complaint to someone more senior than their manager.56

In the EEOC press release concerning the Burger King case, Lynn Bruner expressed her concern for vulnerable adolescents in the restaurant industry. She commented, “We as a society fail when teenagers—as part of their first employment experience—are subjected to graphic language, inappropriate touching, and requests for sexual favors by the very adults who are supposed to make sure they’re safe.”57 She further explained that sexual harassment may be a particular problem in the restaurant industry because restaurants often hire young, inexperienced workers.58 High employee turnover contributes to the problem, presumably because of monitoring difficulties and the need to train new employees continually.59 Bruner also suggested that “restaurants often try to create an ‘entertainment atmosphere’ that can cloud the rules for appropriate conduct in the workplace.”60

3. Youth Sexual Harassment Generally

San Francisco Legal Aid Society-Employment Law Center staff attorney David Pogrel agreed that teenagers face unique risks, stating:

Teenagers on the job are often seen as fungible. A lot of employers don’t treat them with the same respect [as adult workers] . . . . I think that sort of runs over into a whole myriad of rights that are often violated on behalf of young workers—not getting paid correctly and sexual harassment among young women.61

56. EEOC Regional Attorney William R. Tamayo explained that many young workers may not know their rights. Press Release, EEOC, EEOC Settles Sex Harassment Suit with Fresno Chain Uncle Harry’s Bagels (Mar. 13, 2003) (on file with author). On March 13, 2003, the EEOC settled a sexual harassment case against a Fresno, California-based chain, Uncle Harry’s Bagels. Id. A store manager allegedly sexually harassed several teenaged workers, among other employees. Id. Uncle Harry’s agreed to pay $150,000 to the six female victims. Id. According to Sanya Hill Maxim, teens may not recognize sexual harassment and may not know what to do when they experience it. Oreskovic, supra note 46, at 1.


58. Id.

59. Sanchez, supra note 47, at 8.


61. Oreskovic, supra note 46, at 1.
Adolescents face this treatment at a time in their development when they are learning new interpersonal skills and how to function in a workplace environment. They may not yet understand the boundaries between what may be appropriate outside the workplace but illegal within it.62

Sexual harassers prey upon teens in other workplace sectors, not just restaurants. In 1999, the EEOC filed suit against Footaction USA. A manager in his late thirties allegedly sexually harassed a teenaged employee. The manager, co-workers, and customers subjected the teenager to sexual jokes, propositions, and threats, culminating in a physical assault. San Francisco EEOC District Director Susan L. McDuffie commented:

Several incidents of sexual harassment of minors have been brought to our attention this year . . . . Sexual harassment—at what is very often the teenaged employee’s first job—can have a devastating psychological effect, causing the victim to feel shame, to change the way she dresses, to drop out of school, and to be afraid to tell anyone what is happening. We hope this suit will send two messages. Employers must have zero tolerance for sexual harassment. In addition, teens should know they have a right to report unwelcome, offensive sexual conduct whether by a customer, co-worker, or supervisor, and that if they come to the EEOC, we will take action.63

In this passage, McDuffie raised several important points. First, she acknowledged the phenomenon of teen harassment. Second, she briefly explored sexual harassment’s devastating effects on adolescents and mentioned their fear of reporting harassment. Third, she called for zero tolerance by employers in order to protect these teens. Fourth, she emphasized the need for teens to know their rights and that the EEOC responds to complaints. This fourth point supports the notion that some teens do not know their rights and fail to understand that an agency exists to help them.

4. Empirical Research Needed

This anecdotal evidence highlights the need for immediate, comprehensive empirical research in this field. We need to know the prevalence and scope of the sexual harassment of working adolescents.

62. Id.; see also Drobac, supra note 3, at 14–16.
63. Press Release, EEOC, EEOC Sues Footaction USA for Sexual Harassment of Teen Employee (Sept. 29, 1999) (on file with author).
Adolescent Sex & the Workplace

We know, for example, that eighty-three percent of girls and seventy-nine percent of boys report experiencing sexual harassment at school. In the workplace, is sexual harassment of teenagers as pervasive as sexual harassment of adults? Is it even more pervasive? We must determine whether the presence of adolescent workers in an industry correlates with high numbers of sexual harassment charges. Specifically, do harassers target teens rather than adult workers? An investigation into adolescent harassment will lead to other, more sophisticated questions. For example, if sexual harassment of teenagers is pervasive, do they tolerate harassment and, if so, why? Are they ignorant of their rights? Are they immobilized by shame? Is it shame or do they fear economic retaliation, or even physical violence? The American Academy of Child & Adolescent Psychiatry (AACAP) suggested in a policy statement: “It is common for children and adolescents to conceal these offenses [sexual harassment] because they feel afraid, ashamed, vulnerable, and humiliated. They may actually believe their own behavior may have precipitated the sexual harassment. These incidents are often not revealed for many years, if ever.”

The conduct of Sara, whose story begins this Article, supports this statement. She concealed her manager’s abuse and her resulting trauma from her parents. She later reported feeling humiliation and shame. Had the manager not continued to telephone her from prison, Sara’s parents might never have discovered the cause of her plummeting grades and bizarre behavior.

The AACAP statement raises other questions. To what extent do teenagers fail to report harassment? If teenagers are not complaining, do they have unique coping techniques? Do they “consent” to sexual activity and then complain? Is Sara’s case unique? Answers to these questions will enable jurists, educators, healthcare professionals, and

67. Sara Doe Complaint, supra note 1, at 10.
employers to address problems through training, counseling, and legal reform.

While the answers to these questions are important, we can take the anecdotal and EEOC statistical evidence to explore how the law currently treats or fails to address the sexual harassment of adolescents. Additionally, we can explore the public policy reasons for addressing teen sexual harassment. For, as one news reporter noted: “From a legal standpoint, the fast-food industry is liable for alleged incidents. But this is a societal problem. At its most basic level it is an issue of respect. What a sad commentary on American society that so little respect is afforded young women.”

If society truly affords our children—transitioning into adults—so little respect, if employers treat adolescent workers as fungible, then public policy and the law should address the situation.

II. ADOLESCENT “CONSENT” TO SEX—HISTORIC AND CURRENT LEGAL TREATMENT

The criminal and civil legal systems deal with adolescent “consent” to sex in different ways. Statutory rape laws dominate the criminal field. The civil system addresses such consent through state tort laws, Title VII, state FEPS, and other statutory responses. By comparing the varied approaches, one sees the gross legal inconsistencies.

A. Statutory Rape Laws

Historically, statutory rape laws defined “the age of consent” as a girl’s age at which her consent to sexual intercourse earned legal significance and insulated the male participant from criminal prosecution. During the nineteenth century, states raised the age of consent from ten to as high as twenty-one. Currently, all fifty states prohibit the sexual predation of minor females by adults. As late as

---

69. Sanchez, supra note 47, at 8. This reporter’s assertion regarding employer liability is inaccurate. Employers become liable only when the complainant proves that the harassment and damages occurred. Moreover, liability may not attach if a teen “consents.”


71. Id. at 24.

Adolescent Sex & the Workplace

1994, only thirty-five states had gender neutral laws protecting both male and female minors.73 Now, all fifty states protect both sexes.74

In 1997, Charles A. Phipps surveyed state sex crime laws and found that most states distinguish among sex crimes against children by the severity of the offense and the age of the child.75 He concluded that most states classify crimes against children under thirteen or fourteen as the most serious.76 The least serious were non-forcible sex crimes with older children.77 An age difference of at least two to five years was an element of a sex crime in most states.78

Because of the complexity of these laws, I found it difficult to pinpoint a definitive “age of consent.” Some states set a baseline age of consent but then increase the age when a case involves special facts, such as an adult in a position of trust or authority, a relative, or a school employee. For example, four states set the age at fourteen but increase it to sixteen under special circumstances.79 Twenty states increase the age from their respective bases to eighteen under special circumstances.80 If

75. Phipps, supra note 72, at 55–62; see also Charles A. Phipps, Misdirected Reform: On Regulating Consensual Sexual Activity Between Teenagers, 12 CORNELL J.L. & PUB. POL’Y 373 app. A (2003). I disagree with some of Mr. Phipps’ conclusions regarding the age of consent in his more recent work. I have noted my own conclusions concerning the age of consent in App. A, infra pp. 546–73.
76. Phipps, supra note 72, at 57. Phipps also discussed the Model Penal Code’s treatment of sex crimes and highlighted its gendered approach. Id. at 18–19. Moreover, the Model Penal Code sets the age of consent at ten—an age no longer adopted by any state. Phipps explains, “The drafters were worried about the seductive powers of adolescents as well as the application of a rule of strict liability and they wanted to draw a clear line for the most serious offense.” Id. at 19 (citing MODEL PENAL CODE § 213.1 cmt. 6 at 324 (1985)).
77. Id. at 59–60. The age of consent for sexual contact, as opposed to penetration, was lower in most states. Seventeen states set the age of consent for sexual contact at fifteen or below. Another twenty-two set it at sixteen. Id. at 62.
78. See id. at 62 & n.252; Oberman, supra note 70, at 32. Some states set age minimums for culpability. Phipps, supra note 72, at 62–63 & n.248 (citing ALASKA STAT. § 11.41.434(a)(1) (Michie 1996)).
79. See App. A, infra pp. 552, 557, 560, 568 (summarizing the laws for Hawaii, Maryland, Mississippi, and South Carolina).
80. See App. A, infra pp. 546–73 (summarizing the laws for Alaska (16), Arkansas (16), Colorado (15), Connecticut (16), Delaware (16), Florida (16), Iowa (14), Illinois (17), Maine (14), Michigan (16), Minnesota (16), New Hampshire (16), New Jersey (16), New Mexico (17), New York (16), North Carolina (16), Ohio (16), Oklahoma (16), Pennsylvania (16), South Carolina (16), Tennessee (16), Utah (16), and Virginia (16)).
one takes the highest age in each state, then twenty-seven states set the age of consent at eighteen, five at seventeen, and eighteen at sixteen. These laws demonstrate that almost half of the states set the age of consent at below the age of majority. Only fourteen percent of the states (seven) set it at the age of majority absent special circumstances.

Many states recognize an enhancement of the offense when a member of the family or another adult in a position of authority commits the offense. Teachers and other school employees, guardians, babysitters, employers and shift supervisors, psychotherapists, and medical professionals are all examples of persons who wield authority over adolescents.

While statutory rape is a strict liability offense, a few states retain now uncommon elements of the offense or unusual defenses. For example, in California, mistake of age, particularly of older victims, constitutes a defense. In Massachusetts, chastity remains an element of the crimes against older children. Thus, if the child was not virginal at the time of the offense, the perpetrator may use that as a complete defense to the crime of statutory rape.

Oklahoma (16), South Dakota (16), Tennessee (13), Utah (16), Vermont (16), and Washington (16).

82. Id.
83. Id.; Phipps, supra note 72, at 66–69; see also Michelle Oberman, Regulating Consensual Sex with Minors: Defining a Role for Statutory Rape, 48 BUFF. L. REV. 703, 767–68 (2000). As Phipps explains, the Model Penal Code creates an offense by one in a position of authority but narrowly defines that position and excludes the teacher-student relationship. Phipps, supra note 72, at 23–24. The 1980 commentary to the Model Penal Code explains, “Coverage of every instance of sexual relations with an employee, student, or other person under one’s supervision would reach too far.” MODEL PENAL CODE § 213.3 cmt. 4, at 389 (1980).
84. See, e.g., Phipps, supra note 72, at 68 n.264; see also Doe v. Estes, 926 F. Supp. 979 (D. Nev. 1996) (finding that school-aged children are particularly vulnerable to sexual abuse by an adult in a position of authority). The Estes court determined that “[s]choolchildren are particularly vulnerable to mistreatment at the hands of adults, especially where those adults are cloaked with the authority of the state.” 926 F. Supp. at 988; see also infra note 315 and accompanying text.
85. Phipps, supra note 72, at 51–52 & n.219.
86. Id. at 70–71 & n.275 (citing MASS. GEN. LAWS ANN. ch. 272, § 4 (West 1990); MISS. CODE ANN. §§ 97-3-67, -5-21 (1994)). In 1998, Mississippi repealed laws requiring that the target be “chaste.” See MISS. CODE ANN. § 97-3-65 (1999).
Adolescent Sex & the Workplace

1. A Case Example: Hernandez v. State

Until 1993 in Texas, the perpetrator could argue capacity to consent to sex based on a promiscuity defense. In *Hernandez v. State*, the court emphasized:

> We do observe that the State’s brief acknowledges “that children aged fourteen to seventeen who have voluntarily become sexually active are, unlike their sexually inactive peers, imputed by the law with the capacity to consent to sexual conduct like an adult.” It then adds, “Stated conversely, the law which imputes to children an incapacity to consent to sexual activity makes a logical exception for those in their mid-teens who have chosen to become sexually active.”

The court did not explain exactly how sex, or more specifically, the choice to have sex, elevates a minor’s legal capacity to that of adult with full maturity, understanding, and appreciation.

Did the law really impute capacity because these minors made a choice? Did the law assume that every teenager who “chooses” to have sex weighs the decision carefully, considering all the potential ramifications? The court’s reasoning flies in the face of our common understanding of teenage passion, sex, and willingness to engage in risky behavior. The chastity and promiscuity defenses attribute full capacity to minors. What about the *Hernandez* facts? The minor’s mother allegedly “sold” her daughter to the accused rapist and other men. Perhaps, the minor’s “choice” not to resist the accused reflected a well-reasoned decision, but I would hardly attribute to her adult status based on those facts.

To follow this reasoning further, let us ask why both the prosecution and the court suggested that the law made a logical exception for sexually active teens. If teen virgins suffer incapacity, shouldn’t the law protect them from their own naïve and misguided choices? Why would the law credit their choice to have sex in the first place? If the law places

---

87. *Hernandez v. State*, 861 S.W.2d 908, 909–10 (Tex. 1993) (finding that § 22.011(d)(1) of the Texas Penal Code permits the accused to raise a statutory promiscuity defense in a case in which the minor’s mother “sold” her to the accused and other men). The Texas legislature deleted this defense when it revised the penal code in 1993. *Id.* at 910 (Miller, J., concurring).
88. 861 S.W. 2d 908 (Tex. 1993).
89. *Id.* at 909 n.1; see also Oberman, supra note 70, at 33.
90. For a discussion of teenage sexuality, see Drobac, supra note 3, at 16–22.
91. *Hernandez*, 861 S.W.2d at 910 (McCormick, J., dissenting).
such responsibility and resulting consequences on fourteen- to seventeen-year-olds, if it affords their perpetrators the ability to invite a promiscuity defense, then it is not logical to keep statutory rape laws on the books for youth over thirteen.

Another interpretation of the Hernandez court’s reasoning is that the law imputed capacity because sex elevates minors to a new level of understanding. One could argue, using this reasoning, that if we simply introduced all of our teenaged children to sex, they would all develop adult wisdom and legal capacity. Unlikely.

The law in Hernandez reflected adult prejudice. Some adult lawmakers in Texas thought that teenaged children who chose to have sex chose an immoral path and were, therefore, promiscuous and bad. They did not reward or protect bad minors in Texas. In fact, they let adults who raped them avoid punishment. The promiscuity defense, which reflected this adult prejudice, died only ten years ago in Texas. Did the prejudice die with it? I doubt it.

2. Federal Rule of Evidence 412

The chastity and promiscuity defenses clearly contradict the provisions of Federal Rule of Evidence 412 and parallel state statutes that specifically prohibit introduction of the victim’s prior sexual history. When courts fail to invoke Rule 412 (or similar state equivalents) to protect the sexual history of minor victims, then any evaluation of a minor’s maturity could conceivably include a discussion of the minor’s sexual maturity and sexual history. This circumstance leaves minors vulnerable to the grossest procedural abuses that legislators intended Rule 412 (and state equivalents) to combat.

92. I thank Professor R. George Wright for clarifying this point.
93. See Hernandez, 861 S.W.2d at 910 (Miller, J., concurring).
94. Phipps, supra note 72, at 70 n.275; see Fed. R. Evid. 412 & advisory committee’s note: But see Barnes v. Barnes, 603 N.E.2d 1337, 1342–43 (Ind. 1992) (holding that in a tort action, the Indiana Rape Shield Statute did not preclude the defendant from introducing evidence of plaintiff’s other sexual activities); Doe by Roe v. Orangeburg County Sch. Dist., 495 S.E.2d 230, 233 (S.C. Ct. App. 1997) (finding that the South Carolina rape shield statute did not apply in civil cases).
96. See, e.g., Susan Estrich, Sex at Work, 43 Stan. L. Rev. 813, 849 (1991) (arguing that exposure of a woman’s sexual history will “lead not only to shame in the courtroom but acquiescence in the workplace”).
Adolescent Sex & the Workplace

Because Sara lived in California, a state that sets the age of consent for penetration at eighteen, the District Attorney successfully prosecuted a statutory rape charge. Her manager could not raise mistake of age as a defense. He knew (or should have known) her age from her work permit, which California required her to obtain from her school. Had Sara lived in twenty-four other states, however, her “consent” would have insulated the manager from prosecution because she was sixteen, the “age of consent” in those other states, by the time the manager actually seduced her. Thus, the state in which she consented and where the District Attorney prosecuted the claim made a difference in her case.

B. Title VII and State Fair Employment Practice Statutes (FEPS)

Title VII prohibits discrimination against any individual “with respect to his compensation, terms, conditions, or privileges of employment because of such individual’s race, color, religion, sex, or national origin.” In Meritor Savings Bank v. Vinson, the Court held that severe or pervasive sexual harassment violates Title VII when it alters the worker’s conditions of employment and creates an abusive working environment. Courts assess the working environment by “looking at all of the circumstances’ including the ‘frequency of the discriminatory conduct; its severity; whether it is physically threatening or humiliating, or a mere offensive utterance; and whether it unreasonably interferes with an employee’s work performance.’”

In order to bring a case of sexual harassment against an employer under Title VII, a plaintiff must show: (1) membership in a protected

98. See People v. Cosio, No. S9-09852 (Santa Cruz County Super. Ct. 1999).
99. Absent the special facts of her case (e.g., the age disparity and Cosio’s managerial position), the number of states that would not protect her increases to thirty-six. See App. A, infra pp. 546–73.
102. Id. at 67. Professor Catharine MacKinnon defined sexual harassment as “the unwanted imposition of sexual requirements in the context of a relationship of unequal power. Central to the concept is the use of power derived from one social sphere to lever benefits or impose deprivations in another.” CATHARINE A. MACKINNON, SEXUAL HARASSMENT OF WORKING WOMEN 1 (1979).
class; (2) unwelcome sexual harassment; (3) harassment based on sex; (4) an effect on the terms or conditions of employment; and (5) direct or indirect employer liability. No claim against individual perpetrators exists under Title VII. However, some state FEPS permit sexual harassment damage claims against individual harassers. In *Faragher v. City of Boca Raton*, the Court emphasized that the “objectionable environment must be both objectively and subjectively offensive, one that a reasonable person would find hostile or abusive, and one that the victim in fact did perceive to be so.” Jurists refer to the objective component as the “reasonableness” standard and to the subjective element as the unwelcomeness requirement. Every state FEPS that similarly prohibits sex discrimination and sexual harassment also makes “unwelcomeness” an element of the prima facie case.

The *Meritor* Court specifically addressed the issue of volition in its discussion of unwelcomeness:

> While the question whether particular conduct was indeed unwelcome presents difficult problems of proof and turns largely on credibility determinations committed to the trier of fact, the District Court in this case erroneously focused on the “voluntariness” of respondent’s participation in the claimed sexual episodes. The correct inquiry is whether respondent by her conduct indicated that the alleged sexual advances were unwelcome, not whether her actual participation in sexual intercourse was voluntary.

Thus, acquiescence to sex is not consent. The plaintiff must prove only that she somehow indicated that the sexual behavior was unwelcome.

---


105. See *Henson v. City of Dundee*, 682 F.2d 897, 903–05 (11th Cir. 1982).


109. *Id.* at 787 (citing *Harris v. Forklift Sys., Inc.*, 510 U.S. 17, 21–22 (1993)).

110. See, e.g., *CAL. GOV’T CODE § 12940*.


490
Adolescent Sex & the Workplace

One can see how the unwelcomeness requirement, absent invocation of Federal Rule of Evidence 412 (or a similar state statute in combination with a FEPS claim), could lead to the trial of the plaintiff’s conduct.113

Title VII and the FEPS make clear that legislators either neglected adolescent workers when they drafted these laws or knowingly created a glaring conflict between state criminal statutory rape laws and civil antidiscrimination laws. In every state, fifteen-year-old workers lack the capacity in a criminal context to consent to sexual intercourse with a twenty-one-year-old supervisor, but these civil rights laws assume that capacity by making no mention of adolescents. How did this blatant inconsistency pass inspection? This Article reviews the public policy motivating these laws,114 but my guess is that legislators simply forgot adolescent workers.115

Can Sara successfully sue for sexual harassment under antidiscrimination laws? The real question is whether she can prove that her manager’s sexual attention was unwelcome. At first glance, she cannot. She “consented” to sex after initially rebuffing the manager’s advances. Does her ultimate “consent” negate the initial rebuff? Probably not.116 However, she continued to write to her lover while he was in jail.117 Not until she learned that he had lied to her and was a registered sex offender did she again declare his attention unwelcome.118

113. See DEBORAH RHODE, SPEAKING OF SEX 102–04 (1997). Professor Rhode explains:
Although recent reforms seek to restrict disclosures of a complainant’s prior sexual experiences, such evidence is still admissible at trial if the judge decides that its value to a defendant substantially outweighs harm to the complainant. Moreover, in pretrial proceedings, attorneys have greater freedom to ask intimate questions, and can often grill victims about their sex lives, birth control practices, and counseling histories. If a plaintiff alleges physical or psychological damage resulting from harassment, opposing attorneys can explore possible alternative causes for her distress—everything from closeted lesbian experiences to intimate marital difficulties. As a result, defendants’ lawyers can discredit or deter a harassment complaint with harassing tactics of their own.
Id. at 102.

114. See infra Part III.


118. See, e.g., Sara Doe Complaint, supra note 1, at 9.
C. State Personal Injury and Other Tort Claims

Some sexual harassment targets look to common law tort claims, in addition to antidiscrimination laws, for relief. However, most common law intentional tort claims depend upon the plaintiff’s subjective offense and an absence of consent under the maxim *volenti non fit injuria* (“a person is not wronged by that to which he or she consents”). Additionally, consent may trigger defenses to negligence-based tort claims, such as contributory negligence, comparative negligence, and assumption of risk. Finally, many state FEPS provide the exclusive state law remedy for workplace sexual harassment and preempt state common law tort claims. Thus, most “consenting” targets of workplace sexual harassment either cannot avail themselves of common law tort claims or face a trial of their own conduct when they bring common law claims against harassers.

1. Common Law Doctrine and Remedies

Common law was not, historically, without its remedies. Early American civil claims for sexual predation took the form of the writ of
Adolescent Sex & the Workplace

seduction. 124 This claim escaped the consent hazard because the seduced young woman did not sue on her own behalf. Instead, a father sued for his lost honor, now besmirched by his daughter’s damaged reputation, and for his financial losses, tethered to her earning capacity now reduced by pregnancy and motherhood. 125 Consent was irrelevant to the seduction claim. 126

Professor Lea VanderVelde has suggested that two assumptions precluded women from recovering for themselves. She wrote:

First, women lacked the protection against sexual interference afforded by self ownership. Second, the law tacitly assumed that if a woman became involved in a sexual connection she must have consented to the act and all its subsequent consequences. The act of involvement in the event, even as a victim, spoke for the woman, denying her ability to recover against the man. 127

Do these assumptions still apply to working adolescents? As to the first point, one could argue that current laws do not afford minors self-ownership. For example, the law gives parents the right to wages earned by their children. 128 As discussed earlier, children do not enjoy many of the contractual and civil rights enjoyed by adults. 129 They must emancipate themselves to enjoy full legal status. 130
As to the second point, current law, embodied in the “rule of sevens,” explicitly posits the capacity in most teenagers to consent. Under this traditional rule, a minor under age seven cannot give consent, be held liable for negligent conduct, or formulate the requisite mental state to engage in criminal conduct. From seven to fourteen, the law presumes that a minor lacks capacity. From fourteen to twenty-one (now eighteen), a rebuttable presumption declares that minors are competent to consent and responsible for criminal and negligent conduct. Thus, in the context of a civil claim for damages and absent evidence to the contrary, this bright-line rule allows a trier-of-fact to presume that a child over fourteen consents to sexual contact.

Another archaic rule that may relate to modern treatment of sexual harassment victims is the misprision doctrine. Under this rule, a plaintiff risked criminal prosecution if she attempted to sue civilly for her physical injuries before first pressing criminal charges against her rapist or seducer. Thus, the failure to file criminal charges effectively barred a civil suit. Criminal sanctions failed to compensate the victims for the damage to their reputations, medical expenses, and the expense of any child conceived. By funneling the redress for a rape or seduction into the public forum, the government minimized the damage to the victim, maximized the societal harm, and retained its exclusive control over vengeance and retribution. This result has a parallel in modern jurisprudence, as embodied in the new affirmative defense to Title VII claims.

2. An Affirmative Defense to Modern Sexual Harassment Claims

Compare the misprision doctrine to one of Title VII’s new affirmative defenses to sexual harassment charges. In Faragher v. City of Boca
Adolescent Sex & the Workplace

Raton and Burlington Industries, Inc. v. Ellerth,136 the U.S. Supreme Court ruled that a victim’s unreasonable failure to avail herself of an employer’s preventive or corrective procedures insulates the employer from liability for sexual harassment that does not result in a tangible employment detriment.137 Thus, if the employer adopts a complaint procedure that a victim “unreasonably” fails to follow, the Faragher and Ellerth decisions effectively bar her from pursuing a sexual harassment claim.138

The rationale behind the new affirmative defense centers on motivating employers to adopt preventative and corrective procedures.139 The Court’s Faragher and Ellerth decisions encourage victims to complain and employers to cure hostile work environments. A complaint procedure and corrective action, however, cannot remedy the damage already done by a harasser. Moreover, by completely insulating the employer from liability for past harassment, the Court has ultimately charged the cost of this incentive system to the injured victim.140 With this defense, the government (the judiciary) minimizes the harm to the victim, maximizes the harm (discrimination) to society, and gives exclusive control over vengeance and retribution to the employer.

When consent is not an issue and state FEPS do not preempt common law claims, modern plaintiffs can bring a wide variety of tort claims. Such claims include but are not limited to assault, battery, false imprisonment, invasion of privacy, intentional and negligent infliction of

137. The Ellerth Court held:

An employer is subject to vicarious liability to a victimized employee for an actionable hostile environment created by a supervisor with immediate (or successively higher) authority over the employee. When no tangible employment action is taken, a defending employer may raise an affirmative defense to liability or damages, subject to proof by a preponderance of the evidence . . . . The defense comprises two necessary elements: (a) that the employer exercised reasonable care to prevent and correct promptly any sexually harassing behavior, and (b) that the plaintiff employee unreasonably failed to take advantage of any preventive or corrective opportunities provided by the employer or to avoid harm otherwise.

Id. at 765 (citations omitted); see also Faragher v. City of Boca Raton, 524 U.S. 742, 807–08 (1998).

138. Faragher, 524 U.S. at 807–08; Ellerth, 524 U.S. at 765.
139. Faragher, 524 U.S. at 805–06; Ellerth, 524 U.S. at 764; see infra notes 182–83 and accompanying text.

140. See, e.g., Ashton v. Okosun, 266 F. Supp. 2d 399 (D. Md. 2003) (holding that a minor, whose manager touched her on her buttocks and attempted to hug her, had unreasonably failed to avail herself of all complaint procedures when she complained the day after she left work and refused to return after an investigating manager declared her allegations unfounded but offered to transfer her to another shift to avoid the accused).
emotional distress, and loss of consortium or companionship. Other claims focus on the employer’s failure to satisfy a particular duty. Those claims include negligent hiring, negligent supervision, and negligent retention. The minor’s consent operates to weaken, if not extinguish, all of these claims unless some statutory treatment of consent negates the default.141

3.  Consent as a Limiting Factor in Tort Claims

The Restatement (Second) of Torts § 892C offers hope, however, for “consenting” adolescent workers. Subsection (2) states, “If conduct is made criminal in order to protect a certain class of persons irrespective of their consent, the consent of the members of that class to the conduct is not effective to bar a tort action.”142 This guidance suggests that in those states with a high “age of consent” for statutory rape (eighteen), adolescent “consent” should not operate to bar tort recovery.143 The Restatement functions as a guide, but lacks the power of binding precedent.

The Restatement (Second) of Torts § 892A provides meager support for those adolescents in states with lower ages of consent (below eighteen). Subsection (2)(a) specifies that in order to extinguish liability, consent must be “by one who has the capacity to consent.”144 The comment to this subsection provides:

If, however, the one who consents is not capable of appreciating the nature, extent or probable consequences of the conduct, the consent is not effective to bar liability unless the parent, guardian, or other person empowered to consent for the incompetent has given consent, in which case the consent of the authorized person will be effective even though the incompetent does not consent . . . .145

141. One might argue that consent should not extinguish the negligent hiring and supervision claims. However, if the claim that the employer was negligent relates back, for example, to a sexual battery to which the plaintiff consented, then no underlying tort exists and the negligent supervision claim loses its strength.
143. See Wilson v. Tobiasson, 777 P.2d 1379, 1384 (Or. Ct. App. 1989) (holding that a minor’s incapacity to consent to sexual acts under Oregon Revised Statute § 163.315(1) extends to civil cases).
144. RESTATEMENT (SECOND) OF TORTS § 892A(2)(a).
145. Id. § 892A cmt. b.
Adolescent Sex & the Workplace

This exception appears tailored for the mentally challenged, disabled, or immature. According to this caveat, an adolescent’s consent should not be legally binding if the minor could not understand the nature, extent, or probable consequences of the proposed activity. Arguably, a minor, who has never had sex, might not appreciate the consequences of such a choice. Even a non-virginal minor may not appreciate the potential ramifications of sex with a co-worker or supervisor. Scientific information regarding adolescent psychosocial and brain development suggests that adolescents may not have the neurological function and psychosocial ability to formulate legally binding consent.146

Finally, the Restatement (Second) of Torts § 892B addresses some of the special facts in Sara’s case. Subsection (2) states:

If the person consenting to the conduct of another is induced to consent by a substantial mistake concerning the nature of the invasion of his interests or the extent of the harm to be expected from it and the mistake is known to the other or is induced by the other’s misrepresentation, the consent is not effective for the unexpected invasion or harm.147

This provision describes how fraud and misrepresentations can invalidate consent induced by such falsehoods. One might argue that the manager’s profession of love for Sara, his lies about the brain tumor, and his failure to disclose his status as a registered sex offender constituted multiple misrepresentations that vitiated her consent.148 Because she never would have consented to have sex with him had she known the truth, and withdrew her consent once she understood the truth, the fraud paves the way for a sexual harassment suit. Whether a court would follow this non-binding legal guidance remains unclear.149

146. A detailed discussion of adolescent physical and psychosocial development is beyond the scope of this Article. For a complete discussion of adolescent development as it pertains to sexual harassment law, see Drobac, supra note 3. I thank Professor Lois Weithorn who kindly referred me to the MacArthur Juvenile Adjudicative Competence Study that prompted my research of adolescent psychosocial development and its application to sexual harassment law.

147. RESTATEMENT (SECOND) OF TORTS § 892B(2).


149. See Leleux v. United States, 178 F.3d 750, 755 (5th Cir. 1999) (finding that a naval officer’s fraudulent concealment of venereal disease invalidated consent of partner to sexual intercourse). But cf. Food Lion, Inc. v. Capital Cities/ABC, Inc., 194 F.3d 505, 518–19 (4th Cir. 1999) (giving legal effect to consent to an entry and allowing consent as a defense to a claim of trespass, even though it was obtained by misrepresentation or concealed intentions).
D. Conclusions Concerning Criminal and Civil Law

This review of criminal and civil law reveals several truths. First, the law handles adolescent “consent” to sexual conduct inconsistently. The system (criminal or civil), the geographic region (or jurisdiction), and the particular claims alleged all influence the legal treatment of adolescent “consent.” A teenager in California can expect very different treatment than a teenager in Colorado, where the “age of consent” is three years lower.150 Second, common law claims may provide little or no relief to “consenting” teens. Preemption may bar tort claims in employment cases,151 or courts may conclude that a minor appreciated the consequences of her consent to specific conduct. On the other hand, depending upon where the minor works, state criminal law may pave the way for tort recovery via Restatement (Second) of Torts § 892C(2).152 Third, antidiscrimination laws always require that the minor worker find the activity “unwelcome,” and often require the alleged victim to report misconduct.153 Fourth, statutory rape laws draw bright-line rules determining the “age of consent” and denying capacity below that age.154 In sum, no national consensus exists regarding the age of consent or the treatment of adolescent “consent” to a broad variety of adult activities, including sex.

III. PUBLIC POLICY

The question persists why criminal and civil law handle the exact same “consensual” behavior very differently. One response focuses on juridical objectives and public policy motivations. While the goals of the criminal and civil systems overlap in some aspects, they remain distinct in others. The differences in motivating justifications for criminal and civil laws may explain the differing treatment of adolescent “consent.”

A. Criminal Statutory Rape Law

Criminal law prevents harm to society, as well as individuals within society. Our designation of the “State” or the “People” to prosecute

150. See supra notes 70–82 and accompanying text.
151. See supra note 123 and accompanying text.
152. See supra Part II.C.3.
153. See supra notes 109–13 and accompanying text.
154. See supra Part II.A.
Adolescent Sex & the Workplace

perpetrators serves as a constant reminder of that broad utilitarian goal. Another goal of the criminal justice system centers on punishment of the perpetrator who commits the bad act, or actus reus. Punishment serves several subsidiary aims: deterrence (both general and specific), rehabilitation, retribution, and incapacitation. With a strict liability offense such as statutory rape, society demonstrates its concern not with the actor’s guilty mind, or mens rea, but with the harm to our children.

What harm results from “consensual” teen-adult sexual contact? How is this illegal conduct between the adolescent and the adult so different from the same, legal behavior between consenting adults? The question, properly framed, analyzes not so much the sexual conduct but the quality of the consent. The physical acts between teenagers and adults may be the same as those between two adults. What differs between adolescents and adults are the expectations, motivations, and experiential wisdom (or lack thereof) that produce the problematic “consent.” Additionally, because adolescent expectations and motivations differ, and because the nature of the consent differs, the consequences of the exact same behavior also differ for adolescents.

After studying statutory rape cases, Professor Michelle Oberman suggested that “there is at least one important difference between girls and women when it comes to consensual sex: the sexual bargains struck by girls often are so painfully one-sided that it is difficult for adults to understand what prompted the girl to consent.” Professor Oberman categorized a national Westlaw database of statutory rape cases into four groups: consensual relationships, acquaintance rape, stranger rape, and

155. See Phipps, supra note 72, at 27; see, e.g., CAL. GOV’T CODE § 26500 (West 1988) (“The public prosecutor shall attend the courts, and within his or her discretion shall initiate and conduct on behalf of the people all prosecutions for public offenses.”).

156. See Phipps, supra note 72, at 28–29.

157. See Drobac, supra note 3, at 14–16.

158. Harms in addition to juvenile sexual violation justify the statutory rape laws. Phipps explained that in addition to protecting children, statutory rape laws historically protected “the Weaker Sex” and facilitated the prevention of unwanted pregnancy. Phipps, supra note 72, at 34–40; see Michael M. v. Sonoma County, 450 U.S. 464, 470–71 (1981). The Michael M. Court reviewed other possible justifications for the law in validating pregnancy prevention: “Some legislators may have been concerned about preventing teenage pregnancies, others about protecting young females from physical injury or from the loss of ‘chastity,’ and still others about promoting various religious and moral attitudes towards premarital sex.” 450 U.S. at 470. Phipps noted that in 1996, the increasing birth of children to teenagers on public assistance renewed legislative interest in statutory rape laws. Phipps, supra note 72, at 4–5.

159. Oberman, supra note 83, at 714.
overreaching and/or age range of more than ten years. Eighty-one percent of the national group and ninety-one percent of the Illinois cases fell into the overreaching category. Oberman concluded that prosecutors targeted the cases with a greater age gap between the parties because of the greater “risk of significant power disparity between the parties.” Her findings suggest that the public is concerned about coercion and power disparities in sexual relationships involving minors.

The notion that juvenile capacity, and therefore “consent,” differs qualitatively from adult consent engenders passionate debate. This idea, however, dates back hundreds if not thousands of years. Because we have historically deemed children incapable of giving informed consent, the law pertaining to children differs from that concerning adults in everything from contract formation to fundamental civil rights.

If we accept that children cannot give informed consent because they lack capacity, the sexual taking of a child’s body constitutes a theft of the most intimate kind—a rape. This violation, or actus reus, justifies the punishment. As the child approaches the age of consent and maturity, however, society becomes less certain of the disability. We see this uncertainty in the statutes that set lower ages of consent for sexual contact than for sexual penetration.

Children’s rights advocates, who take a self-determinist approach, maintain that children should enjoy the right to make decisions for

160. Id. at 748.
161. Id. at 751.
164. See supra Part I.A (highlighting some of these differences in the review of adolescent rights and legal limitations).
165. See, e.g., Virginia v. Black, 538 U.S. 343 (2003). Justice Thomas stated:
   For instance, there is no scienter requirement for statutory rape. See, e.g., Tenn. St. § 39-13-506; Or. St. § 163.365; Mo. St. § 566.032; Ga. St. § 16-6-3. That is, a person can be arrested, prosecuted, and convicted for having sex with a minor, without the government ever producing any evidence, let alone proving beyond a reasonable doubt, that a minor did not consent. In fact, “[f]or purposes of the child molesting statute . . . consent is irrelevant. The legislature has determined in such cases that children under the age of sixteen (16) cannot, as a matter of law, consent to have sexual acts performed upon them, or consent to engage in a sexual act with someone over the age of sixteen (16).” Warrick v. State, 538 N.E.2d 952, 954 (Ind. 1989) (citing Ind. Code 35-42-4-3). The legislature finds the behavior so reprehensible that the intent is satisfied by the mere act committed by a perpetrator.
   Id. at 397 (Thomas, J., dissenting).
Adolescent Sex & the Workplace

themselves whenever practicable. Some argue that statutory rape laws disregard adolescent capabilities and sexual autonomy. In her recent book *Harmful to Minors: The Perils of Protecting Children from Sex*, Judith Levine argues that adult fears concerning juvenile sexuality combined with the “politics of child protectionism” dominate our governance of children’s sexuality.167 She advocates for “not only protection and schooling in safety but also the entitlement to pleasure.”168 She explains:

There is no distinct moment at which a person is ready to take on adult responsibilities, nor is it self-evident that only those who have reached the age of majority are mature enough to be granted adult privileges . . . .

Legally designating a class of people categorically unable to consent to sexual relations is not the best way to protect children, particularly when “children” include everyone from birth to eighteen. Criminal law, which must draw unambiguous lines, is not the proper place to adjudicate family conflicts over youngsters’ sexuality.169

Levine raises a valid criticism. Her critique of American treatment of child sexuality deserves attention.

In her analysis of the statutory rape laws, Levine recommends the review of Dutch legislation passed in 1990. In Holland, the parliament lowered the age of consent to twelve, but permits adolescents to invoke a consent age of sixteen if those teens conclude that their partners coerced or exploited them. The primary power to decide rests with the


Children should have the right to decide the matters which affect them most directly. This is the basic right upon which all others depend. Children are now treated as the private property of their parents on the assumption that it is the parents’ right and responsibility to control the life of the child. The achievement of children’s rights, however, would reduce the need for this control and bring about an end to the double standard of morals and behavior for adults and children.

Id. See generally RICHARD E. FARSON, BIRTHRIGHTS (1974).

167. JUDITH LEVINE, HARMFUL TO MINORS: THE PERILS OF PROTECTING CHILDREN FROM SEX xxxv (2002).

168. Id.

169. Id. at 88.
adolescent. Parents can preempt their children only if those parents persuade the Council for the Protection of Children that the parents better represent the child’s best interests.\footnote{Id. at 89. Some observers might argue that a child who is too young to have sex is too young to decide whether to invoke the higher age of consent. I thank Professor Florence Roisman for her poignant comment that Holland leaves a critical decision to those youth who, by their own admission, may lack capacity.}

Professor Michelle Oberman also offered a revised version of current statutory rape laws to account for adolescent autonomy while ensuring the protection of children.\footnote{See Oberman, supra note 83, at 778.} Oberman opined:

[O]ne of the most powerful reasons for enforcing statutory rape laws is to set normative parameters around sex so that both boys and girls will learn to honor their own and others’ sexual autonomy . . . .

Thus, to the extent that the law ignores the learning curve at work in adolescent sexual encounters, it may be too harsh. But the failure to condemn “mistakes” involving nonvoluntary sex with an underage partner is equally pernicious. Lenience in such cases only encourages girls to internalize a sexual script that fuses dominance and exploitation with sexual gratification. As such, it vitiates the promise and obligation of statutory rape laws: to make the world safe for boys and girls who are coming of age.\footnote{Id. at 777.}

Rather than lowering the age of consent, Oberman recommended setting the minimum age for consent at no lower than sixteen and asking victims to direct whether the court should suspend the sentence of a first time offender.\footnote{Id. at 778. This scheme avoids the flaw raised in supra note 170: that we burden the child with the decision of whether to prosecute. Under Oberman’s plan, the prosecutor retains prosecutorial discretion. The victim simply makes a plea for a suspended sentence. Oberman also allowed for exceptions to the rule for conduct that appeared obviously coercive or “offensive to societal mores.” Id. at 779. But see Phipps, supra note 75, at 419–21 (rejecting many of Oberman’s conclusions and her final proposal and offering alternate reforms). In his 2003 article, Phipps challenged many of Oberman’s conclusions regarding the prevalence of statutory rape and statutory rape laws generally. However, Phipps emphasized Obereman’s conclusions pertaining to cases involving only adolescents—not adults. See id. at 411. Moreover, Phipps mischaracterized Oberman’s proposal to have victims cooperate with the prosecutions of statutory rape. Id. at 415. Oberman stated, “It strikes me that mandating victim cooperation would be problematic for several reasons” and then discussed those reasons. Oberman, supra note 83, at 779–82. Finally, it appears that Phipps misunderstood Oberman’s plan for reform, suggesting that minor “victims” would direct criminal prosecutions. Phipps, supra note 75, at 420–21. Oberman’s scheme contemplates a role for}
Adolescent Sex & the Workplace

completed a specified course of action, such as mandatory counseling and/or a sex education class, the court could dismiss the charges. If the victim declined to offer an opinion, the prosecutor, not the parents, could make the decision of whether to request a suspended sentence. Finally, Oberman suggested a set of statutorily defined exceptions that would indicate “per se violations.” Statutory law would direct cases with facts suggesting coercion or behavior normatively reprehensible to full prosecution rather than suspended sentence. In the list of exceptions, Oberman included cases involving incest, abuse of authority, repeat offenders, and multiple offenders.

Oberman’s envisioned changes to U.S. statutory rape law acknowledge several public policy goals. First, they recognize society’s determination that underage sex hurts children and that the law should punish and deter it. Second, her plan provides for full prosecution when the perpetrator uses a position of authority or manipulates a clear power disparity to coerce sexual conduct. Third, her proposals reflect an awareness of the developing autonomy and capacity of teenagers to make decisions regarding sexual activity. They attempt to balance a protectionist, nurturance approach with the self-determinist philosophy that would afford adolescents more autonomy as their developing capacity permitted. Under Oberman’s plan, the sexual conduct remains illegal, but older adolescents with sufficient life experience may declare themselves ready for the activity and request the suspended sentence.

B. Civil Antidiscrimination Law

Like criminal law, antidiscrimination law addresses harms to individuals, as well as to society. Professor Catharine MacKinnon described the social harm of sexual harassment: “Sexual harassment exemplifies and promotes employment practices which disadvantage the minor but ultimately, as noted above, leaves discretion to the prosecutor. Oberman, supra note 83, at 778–79.

175. Id. at 778–79.
176. Id. at 779.
177. See id.
178. Id.
179. For a more in-depth analysis of a mixed approach to adolescent rights, see generally FRANKLIN ZIMRING, THE CHANGING LEGAL WORLD OF ADOLESCENCE (1982).
women in work (especially occupational segregation) and sexual practices which intimately degrade and objectify women.”

MacKinnon focused on the subordination of women. Professor Deborah Rhode described the harms more concretely:

For individual victims, harassment often results in economic and psychological injuries, including job dismissals, transfers, coworker hostility, anxiety, depression, and other stress-related conditions. For women as a group, harassment perpetuates sexist stereotypes and discourages gender integration of male-dominated workplaces. For employers and society as a whole, the price includes decreased productivity and increased job turnover. The estimated cost of harassment for a Fortune 500 company averages $8 million a year.

In this passage, Rhode untangled the consequential threads of sexual harassment for the individual target, for women, and for society.

Antidiscrimination laws respond to each level of injury. “For example, Title VII is designed to encourage the creation of antiharassment policies and effective grievance mechanisms.”

“Although Title VII seeks ‘to make persons whole for injuries suffered on account of unlawful employment discrimination,’ . . . its ‘primary objective,’ like that of any statute meant to influence primary conduct, is not to provide redress but to avoid harm.” The U.S. Supreme Court noted that Title VII compensates individuals for their damages but emphasized Title VII’s prevention goal and deterrent effect.

If Title VII serves primarily to prevent abusive sex-based conduct, then legal tolerance of the consent defense as applied to adolescents interferes with Title VII’s deterrence effect. In order to avoid Title VII’s purview, sexual predators will manipulate their vulnerable adolescent targets to “consent.” Without an incentive to prevent “consensual” adolescent sexual exploitation, employers will not create effective policies or warn adolescent workers about coercive or subtly manipulative managers. Obviously, this concern over prevention premises the undesirability of all adolescent workplace sexual conduct.

180. MacKinnon, supra note 102, at 7.
184. Similarly, the admission of consent evidence in an evaluation of fault and damages interferes with Title VII’s deterrent effect by redistributing responsibility for the conduct.
Adolescent Sex & the Workplace

including “consensual” behavior. Because almost seventy percent of adults think fourteen- to sixteen-year-olds should not engage in sexual relations, and because seven states define the age of consent unequivocally at eighteen, this presumption concerning the undesirability of “consensual” adolescent sexual conduct in the workplace seems reasonable.

C. Tort Law

Professor Kenneth Abraham described five functions served by the imposition of civil liability for accidental and intentional injuries. First, many people expect civil tort liability to compensate individuals for their injuries. Abraham reasoned that compensation alone is not really a goal of the tort system, or anyone injured could obtain compensation. Instead, Abraham focused on tortious conduct and other goals, including deterrence. Like antidiscrimination laws, tort liability serves to deter bad acts and prevent future harm. Abraham also mentioned a third function, loss distribution. Rather than make the plaintiff bear the burden of any loss, tort law provides for a broader distribution. The tort system contemplates that insurers will pay damages incurred by policy holders and that those prospective defendants who can increase the costs of their goods and services to pay future judgments will do so.

Fourth, Abraham suggested that corrective justice warrants civil liability in that it restores the moral balance between parties. For

185. See supra note 34.
186. The number of states increases to twenty-seven if a case includes special facts, such as a significant age difference between the parties or an adult in a position of trust or authority. See App. A, infra pp. 546–73.
187. ABRAHAM, supra note 120, at 14–19.
188. See id. at 17.
189. Id. at 17–18.
190. Id. at 15. Abraham ultimately decided that tort law “does not serve any single goal . . . [but] performs a ‘mixed’ set of functions.” Id. at 19.
191. Abraham modified this goal by calling it “optimal deterrence.” Id. at 19. He explained that, under this justification, liability should deter only excessively risky conduct. Some conduct is not risky enough to justify the deterrence costs that people might incur. See id. at 15.
192. Id. at 17.
193. Id. at 16.
194. Id. at 16–17.
195. Id. at 14–15.
example, when one person intentionally injures another, basic fairness justifies the imposition of liability for damages on the bad actor. Corrective justice seems less appropriate when someone other than the actor, such as a corporate employer, must pay the damage award. In that situation, the disconnect between the intentional harm and the remedial compensation lessens the moral justification. The loss distribution function better justifies the imposition of tort liability in such a case—especially when the tortfeasor lacks deep pockets to pay for the damage caused.

Finally, Abraham reviewed the redress of social grievances through the tort system. Again, much like the function of antidiscrimination law, tort law arguably “is a populist mechanism that permits ordinary people to put authority on trial.” This conception of tort law envisions the amelioration of problems that affect society more broadly.

Taken together, the goals of the tort system harmonize nicely with the desire to protect adolescent workers. Arguably, we want to deter sexual behavior that could harm these workers. By making employers liable for the acts of their agents, we distribute the losses associated with injuries to teens who cannot easily cover those losses. Additionally, tort liability supports corrective justice by restoring the moral balance between employers and their agents, and adolescent workers who may not have the experience or wisdom to recognize manipulative sexual advances. As “a populist mechanism that permits ordinary people to put authority on trial,” tort law allows adolescents to challenge supervisors, managers, and employers. When tort law affords adolescent “consent” legal significance, it fails those young workers and thwarts the goals of the system.

D. Conclusions

This review of the law and public policy demonstrates that statutory rape law, antidiscrimination law, and tort law share many similar functions and public policy goals. Each attempts to deter and prevent antisocial, harmful behavior. The sexual exploitation of minors conceivably falls into the set of antisocial behaviors under criminal, tort, and antidiscrimination law. Each system holds actors who cause harm

196. Id. at 15.
197. Id. at 16–17.
198. Id. at 19.
199. Id.
Adolescent Sex & the Workplace

responsible, by punishing them with incarceration or by awarding damages against them. Each system operates to protect potential victims from harm through the deterrence and prevention mechanisms. The only function that the criminal system does not share with civil law is the redistribution of the costs of harm suffered. Typically, criminal laws do not compensate the victims for their losses. Both antidiscrimination and tort laws provide for such compensation, with damage awards against either the tortfeasor or another responsible party, such as an employer or insurer.

The obvious next question is whether the civil compensation function explains why the criminal and civil systems treat adolescent “consent” so differently. The answer is not immediately apparent. If minors lack capacity to consent in the criminal arena, why might civil courts consider such consent in the redistribution of the costs of injury—especially at the minor’s expense? Courts that resolve civil adolescent sexual harassment cases must deal with the contradictions and conflicts between civil and criminal laws. Not surprisingly, they deal with adolescent “consent” in inconsistent ways.

IV. CASE LAW

Even with five million adolescents in the United States workforce, no Title VII case decisions discuss adolescent “consent.” Common sense dictates that some of these adolescent workers “consented” to sexual contact. Sara’s situation proves that such cases exist. Where are the Title VII decisions then? Do these cases all settle before trial? Are they decided without opinions and never appealed? The unwelcomeness requirement may deter many youth, but surely not all.

Several factors may explain the dearth of Title VII cases. First, practical considerations may prevent prosecution of these cases. Concerned that adolescent “consent” may constitute a complete defense

200. Of course, any suit would have to be brought by the parents or guardian because minors lack the capacity to sue in federal and state courts. How ironic that United States’ antidiscrimination laws attribute to minors the ability to consent to sexual relations but not the capacity to seek judicial relief in court for their own injuries.

201. An informal poll by Teen People magazine reported forty-seven percent of teenage girls said they had been touched against their will at work. Glenn Burkins, A Special News Report About Life on the Job—And Trends Taking Shape There, WALL ST. J., Apr. 20, 1999, at A1. I find it incredible that they all protested or otherwise indicated that they found this contact unwelcome.

202. These cases may settle quickly because the media impact of a public sexual harassment trial involving a minor victim might engender much more expense and ill will than a quick settlement.
and disqualify minors from suit, lawyers may decline to take these cases on a contingent fee basis. Minors and their parents may lack the financial resources to pay an attorney an hourly fee. Second, as the AACAP Policy Statement suggested, adolescents may think that they somehow caused the problem or “attracted” the inappropriate attention.203 Humiliated and ashamed, they may decide not to report the abusive behavior.204 Third, many adolescents may not even realize that they have been sexually harassed or that they can bring a claim for damages.205 A child that does not appreciate the nature of the conduct or its consequences may not appreciate its illegal nature. Empirical research can answer whether these cases abound and why adolescents refrain from filing suit. These listed reasons are just a few of the plausible explanations for the nullity of Title VII cases on the issue of consent.

A review of tort and statutory adolescent sexual harassment “consent” cases from across the nation reveals that they fall into four different groupings with several that fit into two or more categories.206 The largest

203. See Am. Acad. of Child & Adolescent Psychiatry, supra note 65; see also Doe v. Estes, 926 F. Supp. 979, 988 (D. Nev. 1996). The Estes court explained:

Children are often reluctant to report invasions of their bodily integrity. They may fear reprisals by their attackers, they may harbor doubts that their attackers’ fellow grownups will display sympathy or willingly credit their accounts, and they all too frequently are paralyzed by the shame that attends subjection to sexual abuse.

926 F. Supp. at 988.

204. See, e.g., Doe v. Taylor Indep. Sch. Dist., 15 F.3d 443, 449 (5th Cir. 1994) (noting that “Doe explained that she had kept the matter [sexual conduct with a teacher] a secret because she feared the repercussions of disclosure”); Leach v. Evansville-Vanderburgh Sch. Corp., No. EV98-0196 C-Y/H, 2000 WL 33309376, at *2–*3 (S.D. Ind. May 30, 2000) (finding that a student failed to report sexual harassment because she felt ashamed, was afraid to tell her mother for fear of upsetting her, and was afraid that no one would believe her that a teacher was sexually fondling her).

205. See supra note 56. Minors may also miss the applicable statute of limitations for filing their claims. See, e.g., Ashton v. Okosun, 266 F. Supp. 2d 399, 404 (D. Md. 2003) (refusing to toll the 300-day statute of limitations for a minor to bring suit under Title VII).


I do not consider cases decided before 1945 because of the very different prevailing sexual norms and changes in law in the succeeding years. See, e.g., Barton v. Bee Line, Inc., 265 N.Y.S. 284, 285 (N.Y. App. Div. 1933) (holding that consent negates civil claim for damages resulting from a rape); Braun v. Heidrich, 241 N.W. 599, 601 (N.D. 1932) (finding that minor guilty of fornication may not
Adolescent Sex & the Workplace

category (seventeen cases) deals not with sexual harassment per se, but with state and common law civil claims based upon alleged inappropriate sexual conduct with a minor.207 The next largest group (four cases) involves sexual harassment but not in employment. It comprises those cases brought under Title IX of the Education Amendments of 1972208 because the harassment happened at school or the perpetrator was a teacher or staff member.209 In several cases (three), plaintiffs used § 1983210 when the perpetrator was a public school

---


209. Mary M. v. N. Lawrence Cmty. Sch. Corp., 131 F.3d 1220, 1227–28 (7th Cir. 1997) (deciding both Title IX and § 1983 claims), cert. denied, 524 U.S. 952 (1998); Bostic, 2003 WL 723262, at *4–8 (disposing of Title IX and state common law claims); Hackett, 238 F. Supp. 2d 1330, 1346, 1356–60 (resolving Title IX, § 1983, and Georgia common law claims); Benefield v. Bd. of Trs. of the Univ. of Ala., 214 F. Supp. 2d 1212, 1218 (N.D. Ala. 2002) (deciding only a Title IX claim).

employee.\textsuperscript{211} Under § 1983, a student sued for a violation of her Fourteenth Amendment substantive due process liberty interest right to bodily integrity.\textsuperscript{212} I found only one state FEPS employment case for an incident that occurred at a restaurant.\textsuperscript{213}

\textbf{A. FEPS: Doe v. Mama Taori’s Premium Pizza}

In \textit{Doe v. Mama Taori’s Premium Pizza},\textsuperscript{214} a sixteen-year-old male and his parents brought suit under the Tennessee FEPS for sexual harassment against a male adult co-worker and the restaurant owner.\textsuperscript{215} Doe claimed he engaged in sexual acts in the restaurant bathroom with thirty-two-year-old Christopher Abson, who had given him “a ‘marijuana cigarette [that] contained a “knock out drug” that caused [him] . . . to become incapacitated.’”\textsuperscript{216} The parents alleged that Abson had two prior rape convictions, one for raping a child.\textsuperscript{217} Abson pled guilty to two counts of the statutory rape of Doe.\textsuperscript{218} Via an interlocutory appeal in the civil case, the plaintiffs challenged the trial court’s refusal to strike the restaurant owner’s affirmative defenses based on the minor’s consent and comparative fault.\textsuperscript{219}

\textbf{1. Consent—A Civil Defense?}

Doe and his parents first argued that consent should not constitute a defense to a civil action for damages when it fails as a defense to a

\begin{itemize}
\item \textsuperscript{211} Mary M., 131 F.3d at 1220; Doe v. Taylor Indep. Sch. Dist., 15 F.3d 443, 449–50 (5th Cir. 1994) (resolving only a § 1983 claim); \textit{Hackett}, 238 F. Supp. 2d at 1330.
\item \textsuperscript{212} Taylor, 15 F.3d at 451.
\item \textsuperscript{215} Doe and his parents made claims under the state FEPS, the Tennessee Human Rights Act, as well as intentional tort and negligence claims. \textit{Id.} at *2. Same-sex harassment based on sex violates Title VII. \textit{Oncale v. Sundowner Offshore Servs., Inc.}, 523 U.S. 75, 78 (1998). Most states have similarly interpreted their antidiscrimination FEPS.
\item \textsuperscript{216} \textit{Mama Taori’s}, 2001 WL 327906, at *2.
\item \textsuperscript{217} \textit{Id.} at *1.
\item \textsuperscript{218} \textit{Id.} at *2. Mr. Abson also pled guilty to contributing to the delinquency of a minor. \textit{Id.}
\item \textsuperscript{219} See \textit{id.}
\end{itemize}
Adolescent Sex & the Workplace

criminal statutory rape charge. The court disagreed, classifying the sexual behavior as a battery. Referring to well-established principles of common law, the court held that a plaintiff who consents cannot later complain of the behavior. The court acknowledged that consent lacks defensive significance “if (1) the person giving consent lacked the necessary capacity, (2) the consent was coerced, (3) the person giving the consent was mistaken about the nature and quality of the act, or (4) the nature of the act was such that no person could consent to it.” The court also explained that “[i]ncapacity to give consent may arise from age, intoxication or mental incompetence.” However, the court declined to rule, as a matter of law, that Doe lacked capacity simply because he was under eighteen. Doe’s capacity and the quality of his consent remained triable issues of fact.

In discussing the capacity of minors to consent, the court relied upon a medical consent case, Cardwell v. Bechtol. Quoting Cardwell, the court found “that maturity is now reached at earlier stages of growth than at the time the common law recognized the age of majority at 21 years.” The court listed the Cardwell factors that determine a mature minor’s capacity: “age, ability, experience, education, training, degree of maturity or judgment, [and] . . . the minor’s conduct and demeanor at the time of the incident.” The Mama Taori’s court avoided a detailed evaluation of the nuanced indicia of maturity and training, however. The Mama Taori’s court endorsed the “rule of sevens,” used in Cardwell, as guidance for determining whether a minor has the capacity to give consent. Thus, the Mama Taori’s court found Doe presumptively capable of giving consent, and thereby presumptively insulated his co-worker (and the restaurant owner) from liability.

220. Id. at *4. The age of consent in Tennessee is eighteen when the perpetrator is at least four years older than the target who is at least thirteen. See TENN. CODE ANN. § 39-13-506 (2003).
221. See Mama Taori’s, 2001 WL 327906, at *4.
222. Id.
223. Id. (citing RESTATEMENT (SECOND) OF TORTS §§ 892A(2), 892B).
224. Id. (citing Cardwell v. Bechtol, 724 S.W.2d 739, 746 (Tenn. 1987)).
225. See id. at *8.
226. 724 S.W.2d 739 (Tenn. 1987).
228. Mama Taori’s, 2001 WL 327906, at *5 (citing Cardwell, 724 S.W.2d at 748) (internal quotations omitted).
229. Id.; see supra notes 131–33 and accompanying text.
The Mama Taori’s court bolstered its conclusion that minors develop the capacity to consent before age eighteen by reviewing the post-Cardwell “mature minors” doctrine and the Tennessee law that affords minors legal decision-making capacity in a variety of contexts. First, the court discussed consent to medical treatment, specifically abortion, birth control information and supplies, and treatment for drug abuse.231 With respect to a minor’s capacity to consent to abortion, the court ignored the fact that a minor must typically have a parent’s permission or be adjudged competent by a court to make that decision before obtaining an abortion.232 Surely, in referencing abortion, the Mama Taori’s court was not suggesting that a minor should seek judicial by-pass for a capacity determination prior to consenting to sexual relations at the workplace.

With respect to contraception and drug abuse treatment, the Mama Taori’s court failed to address the public policy rationale for those provisions. Laws do permit minors to obtain contraceptives and seek treatment for drug abuse and sexually transmitted diseases, but not because society considers those minors competent to make medical decisions without parental involvement.233 To the contrary, we know that many of those minors need medical treatment because they made immature and uninformed decisions to have unprotected sex, thus demonstrating their lack of competence to make well-reasoned decisions regarding sexuality.234 Rather, as a matter of public policy, we have determined that medical assistance is so important, and the potential

231. Id. at *5.


233. See Scott, supra note 32, at 568 & n.80. Scott explained:

No one argues that minors should be deemed adults because they are particularly mature in making decisions in these treatment contexts. Rather, the focus is on the harm of requiring parental consent. The targeted treatments all involve situations in which the traditional assumption—that parents can be counted on to respond to their children’s medical needs in a way that promotes the child’s interest—simply might not hold. For example, some parents may become angry upon learning of their child’s drug use or sexual activity. Moreover, even if most parents would act to promote their children’s welfare, adolescents may be reluctant to get help if they are required to inform their parents about their condition, either because they fear their parents’ reactions or because they do not want to disclose private information. Removing this obstacle encourages adolescents to seek treatment that may be critically important to their health. Of course, society also has an interest in reducing the incidence of sexually transmitted diseases, substance abuse, mental illness, and teenage pregnancy. Together, these social benefits largely explain why lawmakers shift the boundary of childhood for the purpose of encouraging treatment of these conditions.

Id. at 568 (footnote omitted).

negative consequences of no treatment are so devastating, that we will facilitate the medical treatment of minors, despite their immaturity.\(^{235}\)

The Mama Taori’s court also noted that Tennessee law criminalizes sex with teenagers over thirteen only if their partners are more than four years their senior.\(^{236}\) Tennessee’s failure to criminalize sex between two fourteen-year-olds does not necessarily mean that the Tennessee legislature has judged these minors emotionally, psychologically, and physically mature, however. The permissive stance concerning young teens, associated with the required age differential for prosecution, results more likely from the targeted obviation of predation by mature adults, than a judgment that young teens are ready to appreciate fully the ramifications of their sexual behavior. When neither teen has the capacity to consent to sex, why prosecute one (or both)?\(^{237}\) Laws pertaining to underage sex may not deter those youth who lack the capacity to appreciate the ramifications of their sexual behavior. When the threat of death, from AIDS or other sexually transmitted disease, fails as a deterrent, we cannot expect that statutory rape laws will inhibit teen sexual exploration.

In addition, age difference requirements in statutory rape laws say little about the ability of adolescents to consent at the workplace. Even if the Mama Taori’s court properly used this age differential element to demonstrate that a sixteen-year-old has the capacity to consent to sex with peers away from the workplace, one cannot conclude from that determination that the same youth possesses the capacity to consent to sex with a thirty-two-year-old co-worker. The logic simply fails. If anything, the age difference requirement and its associated concern for power disparities are even more relevant at the workplace. At work, a manager or an adult with more work experience enjoys a position of

\(^{235}\) See Oberman, supra note 70, at 47; Scott, supra note 32, at 568.

\(^{236}\) Mama Taori’s, 2001 WL 327906, at *6.

\(^{237}\) See Allstate Ins. Co. v. Patterson, 904 F. Supp. 1270, 1282 (D. Utah 1995). In an insurance coverage case, the Patterson court explained:

If the age of the victim is relevant, then arguably the age of the perpetrator should also be relevant. The reason a child lacks the capacity to consent to sexual activity is because the child cannot fully appreciate the consequences of such activity. But if a child cannot fully appreciate the consequences of sexual activity, that is reason not to hold the child perpetrator to the same standard as an adult . . . . [C]riminal statutes that make it a crime to engage in sexual acts with a minor regardless of the minor’s consent are based on the premise that minors lack the experience necessary to give meaningful consent, yet “courts cannot seek to protect naïve fourteen-year-old[s] . . . on the one hand, while inferring the most degrading and unnatural [intentions] on the other hand.”

greater status, power, or seniority, and can more easily seduce an inexperienced minor.

The Mama Taori’s court equated the capacity of adolescents to consent to teen-adult sex with their ability to engage in a broad variety of activities: lease safety deposit boxes, work part-time, obtain a driver’s license, execute a durable power of attorney for healthcare, and consent to sex with a peer. The court drew completely inapposite analogies. Bright-line age demarcations account crudely for the fact that adolescent capacity varies by individual and, more particularly, by specific situation. To suggest that an adolescent worker has the capacity to consent to oral sex at the workplace with an adult co-worker because Tennessee law allows him to lease a safety deposit box would be laughable if the results of such a conclusion were not so potentially devastating. The court repeatedly failed to recognize the more plausible motivating reasons for permitting minors certain liberties normally reserved for adults. For example, we allow minors to surrender their children for adoption, not because we think that those minors are sufficiently competent to make the adoption decision, but because we consider them incompetent to raise children.

The Mama Taori’s court found that the “mature minors” rule as well as the “rule of sevens” governed this case and presumptively determined that Doe’s consent carried legal weight. In sum, the court placed the burden of proof on Doe, requiring him to overcome the presumption of capacity. The court’s further discussion of the “totality of the circumstances” and the Cardwell factors was irrelevant since the court never engaged in a detailed analysis of Doe’s maturity. Later in this Article, I address the advisability of engaging in such analysis. Suffice it to say here, with a reference back to the Texas Hernandez case, that a serious concern arises. Any sexual harassment case could devolve into a trial of the minor’s maturity (and morality), rather than of the perpetrator’s culpability and the principal’s liability. Do we really want our children on trial as they seek remedy for sexual abuses?

239. Id. (citing TENN. CODE ANN. § 36-1-110(a) (1996)).
240. Id.
241. The court postponed this more detailed analysis for subsequent proceedings. Id. at *11.
242. See infra Part V.A.
243. See supra Part II.A.1.
Adolescent Sex & the Workplace

2. Consent’s Admissibility

The Mama Taori’s court also rejected the plaintiffs’ reliance on criminal law to prevent the jury from even considering Doe’s consent. The court suggested that barring consideration of consent “would permit any victim of statutory rape to recover civil damages notwithstanding the circumstances.” The court listed three reasons for its rejection of plaintiffs’ argument: the courts that have declined to adopt this per se liability rule have recognized (1) that the statutory rape laws do not explicitly create a private right of action for damages; (2) that criminal and civil proceedings have different purposes; and (3) that it is fundamentally unfair to permit a civil litigant to obtain money damages while preventing the trier-of-fact from considering relevant evidence regarding damages and credibility.

In this analysis, the court found first that statutory rape laws create no private civil claim. However, they do enforce a public consensus that adolescents lack the capacity to consent to sexual intercourse and, at certain ages, to sexual contact of any kind. Statutory rape laws also reflect society’s determination that underage sex hurts children. The Mama Taori’s court deflected attention from the relevance of the consent defense by focusing on the availability of civil damages in a private civil claim. The court did not explain how the availability of damages transforms the nature of capacity to consent. More particularly, the court neglected to explain why the lack of a private claim in the criminal statute precludes, in a civil claim, treatment of adolescent “consent” consistent with the criminal law’s approach.

Second, the court noted the different purposes of criminal and civil law with regard to the admissibility of consent. However, our recent review of the public policy goals suggests that criminal and civil law functions are more similar than they are different. Statutory rape laws have a three-fold purpose: to protect children (individuals subordinated because of their immaturity) from sexual predators, to punish

244. Mama Taori’s, 2001 WL 327906, at *7.
245. Id. (citing Beul v. ASSE Int’l, Inc., 233 F.3d 441, 450–51 (7th Cir. 2000); McNamee v. A.J.W., 519 S.E.2d 298, 302 (Ga. Ct. App. 1999)).
246. Id. (citing Cynthia M. v. Rodney E., 279 Cal. Rptr. 94, 97 (Cal. Ct. App. 1991)).
247. Id. (citing LK v. Reed, 631 So. 2d 604, 607 (La. Ct. App. 1994); Doe by Roe v. Orangeburg County Sch., 518 S.E.2d 259, 261 (S.C. 1999)).
248. See supra note 34 and accompanying text.
249. See supra Part III.
perpetrators, and to deter predators.\textsuperscript{250} Civil rights claims such as Title VII and FEPS serve similar purposes: to protect those workers (subordinated because of their weaker status) who experience discrimination (often in the form of sexual predation) and to punish and deter those employers who permit employees (or agents) to prey upon subordinated individuals.\textsuperscript{251} The primary difference relates to the availability of damages, intended to make the person whole in the civil case. The civil system compensates victims for their injuries and influences employers with the threat of a financial penalty.

Thus, the different purposes of the civil and criminal systems, and the second reason for the court’s treatment of Doe’s consent, boils down to the availability of damages to compensate the victim in the civil system. The Mama Taori’s court had referred to this difference, the right to damages, as the first reason for a rejection of a per se liability rule. The second of the court’s reasons collapses into the first and still makes no sense. The court again failed to explain why the lack of a private claim for damages in the criminal statute precludes, in a civil claim, treatment of adolescent “consent” consistent with the criminal law’s approach.

The third reason reveals the court’s motivating concerns: credibility and damages. This reason centers on the alleged unfairness of permitting the recovery of money damages absent a hearing of all relevant evidence, including consent, which may provide guidance on the evaluation of credibility and the calculation of damages. It also highlights several problems with the court’s treatment of adolescent “consent.” If a sixteen-year-old lacks the capacity to consent to teen-adult sex in Tennessee under criminal law, how does this same consent miraculously shed light on either credibility or damages under civil law? If the minor lacks the capacity to consent, then credibility is not relevant as long as the sex occurred. The sex constitutes an offense no matter what the adolescent said at the time of the incident or says later in court. It qualifies as an offense because society (the legislature) has determined that teen-adult sex is offensive and injurious.

Is society misguided? Is teen-adult sex inoffensive? The reality of the Mama Taori’s case is that Abson did not really “hurt” society by having sex with Doe. This case was not a criminal prosecution, so it lacked the underlying criminal public policy concern for “society.” Abson may

\textsuperscript{250} See supra Part III.A.

\textsuperscript{251} Faragher v. City of Boca Raton, 524 U.S. 775, 805–06 (1998); see supra note 183 and accompanying text.
Adolescent Sex & the Workplace

have outraged those adults who learned of the behavior, but “society” did not change. The reality is that Doe suffered the brunt of whatever injury Abson caused. Moreover, society (via the prosecutor) did not sue for its own damages. Nor did society sue on behalf of Doe. His parents did. Here rest the problems.

The credibility determination for the *Mama Taori’s* court concerns not whether the sex occurred, but the severity, and even existence, of consequential damages. The *Mama Taori’s* court wanted to ensure that the jurors had the opportunity to evaluate Doe’s credibility regarding his personal offense and damages. Society’s determination that teen-adult sex is criminally offensive in Tennessee was irrelevant to the *Mama Taori’s* civil court. Society’s determination evaporated, and the court started from a clean civil slate, demanding that Doe prove his personal offense and damages. This civil court equated consent with immunity from damage (*volenti non fit injuria*) and did not equate societal harm with individual harm. The judge and jurors may have felt outraged about the homosexual “rape,” but their offense was not Doe’s, and the judge wanted the jurors to evaluate Doe’s personal offense and damages.

Doe’s consent fostered the judge’s skepticism about Doe’s damages and outrage. Why? Because jurors might find that Doe was morally culpable, an accomplice in an illicit act. The court even noted in the recitation of the facts that, prior to the alleged sexual encounter, when *Mama Taori’s* transferred Abson to another restaurant, Doe sought a transfer to the same location. When Mama Taori’s denied Doe the transfer, he threatened to quit and reapply at the other location.252

These facts are irrelevant to the issue of capacity for several reasons. That Doe may have admired or even adored Abson does not prove that Doe had the cognitive and psychosocial ability to formulate legally significant consent to sex. Second, Tennessee statutory rape law that sets the age of consent at eighteen indicates the legislature’s determination that someone Doe’s age does not have capacity.253 Third, reliance on these facts risks the importation of a defense similar to the chastity and promiscuity defenses. Recall *Hernandez.*254 In *Hernandez*, the court ruled that consent to sex elevates a minor’s capacity to one of legal significance.255 As noted previously, the choice to engage in sex is not a

---

255. *Hernandez*, 861 S.W.2d at 910 n.1.
scientific indicator of maturity and capacity. Neither is the choice to follow a worker to another employment site. These facts paint Doe as a willing accomplice and foreshadow the court’s true concerns—moral culpability and legal responsibility. Ultimately, the court conflated the two and found Doe (potentially) legally responsible because he was (possibly) morally culpable.

3. Developmental Capacity Applied to Mama Taori’s

I distinguish between legal responsibility and moral culpability or blameworthiness. If I accidentally break a glass, I am responsible but not morally culpable because I did not smash the glass purposefully. With the rule of sevens and the infancy defense, we shield young children from legal responsibility for their criminal and negligent behavior because we adjudge them incapable of understanding or avoiding criminal and tortious conduct. We believe them innocent, morally blameless.256

Consider this distinction step by step as it relates to Mama Taori’s. Begin with the notion that no capacity equates with no fault. Graduate to the concept that teenagers have some capacity. I find the term “diminished capacity”257 inappropriate because the word “diminished” carries a negative connotation. Additionally, it suggests that full capacity should exist or may once have existed. Most teenagers suffer not from impairment but from immaturity—a blameless condition and a natural phase of growth. I prefer the term “developing capacity” because of a teenager’s transitional status from childhood to adulthood and his or her


257. Elizabeth S. Scott & Laurence Steinberg, Blaming Youth, 81 TEX. L. REV. 799, 801, 829–34 (2003). In discussing adolescent criminal offenders, Professors Scott and Steinberg explained: The differences that distinguish adolescents from adults are more subtle—mitigating, but not exculpatory. Most obviously, cognitive and psycho-social immaturity undermines youthful decisionmaking in ways that reduce culpability. Moreover, due to their immaturity, adolescents may be more vulnerable to coercive pressures than are adults. Finally because their criminal acts are influenced by normal developmental processes, typical adolescent law breakers are different from fully responsible adults whose crimes are assumed to be the product of bad moral character. Thus, young offenders are less culpable than adults because of their diminished capacity; but they are also appropriately identified with actors who succumb to coercive pressures or who demonstrate that their criminal acts were out of character, and who are less culpable because their responses are those of ordinary persons.

Id. at 829–30.
Adolescent Sex & the Workplace

developing maturity.\textsuperscript{258} What level of fault should we associate with developing capacity? “No fault” hardly seems fair since the teenager has some capacity. Logically, one could equate the quantum of fault with the level of maturing capacity. Thus, we see how a comparative fault scheme would appear attractive to a court attempting to associate fault with capacity. Here, however, is the flaw in the logic.

Full legal capacity means just that—complete capacity. Not diminished capacity. Not developing capacity. Full legal capacity is an all-or-nothing proposition. There is no sliding scale for legal capacity. Our discussion in Part I confirmed that adolescents have not reached that legal threshold. Even in the criminal system, we try adolescents in juvenile court as children or in adult court as adults. We do not try them in adult court as mature children. When one considers such a sliding scale seriously, one sees the fallacy of such an idea. How can we justify holding someone morally culpable and then fully legally responsible, when that person is incapable of manifesting full reasoning and decision-making abilities because of transitioning developmental maturity?

I realize that this stance necessarily leaves intact an inconsistency between the criminal and civil systems. Because of the need to protect society from crimes committed by adolescents, I endorse Professors Elizabeth Scott and Laurence Steinberg’s proposal that the juvenile justice system recognize adolescent “diminished responsibility” due to diminished culpability.\textsuperscript{259} However, I reassert that adolescents—even adolescent criminal offenders—lack full adult legal capacity. Moreover, I do not suggest a “diminished culpability” or “diminished responsibility” parallel for the civil system because my focus is the protection of youth, and their developing capacity, from sexual exploitation by adults. I would still shield adolescents from legal responsibility for their immature choices because adult exploitation causes their injury. The need to protect society (and individual victims) from crimes committed by adolescents, however, justifies the different treatment in the criminal system of adolescent “developing capacity” and the different level of legal responsibility (and culpability) attributed to adolescent criminal offenders.

\textsuperscript{258} I thank Martin Drobac for clarifying this point.

\textsuperscript{259} Elizabeth Cauffman et al., Justice For Juveniles: New Perspectives on Adolescents’ Competence and Culpability, 18 QUINNIPIAC L. REV. 403, 405 (1999); Scott & Steinberg, supra note 257, at 835–36; see also Scott, supra note 32, at 589–96; Laurence Steinberg & Elizabeth Cauffman, The Elephant in the Courtroom: A Developmental Perspective on the Adjudication of Youthful Offenders, 6 VA. J. SOC. POL’Y & L. 389, 399, 405–06 (1999).
Consider another example using a younger child and compare it to the Mama Taori’s case. A stranger in a car offers candy to a six-year-old on the street. If the child consents to enter the car, do we hold the child legally responsible for his damages that result? No. First, the child lacks the experience and knowledge to comprehend that he should not follow the stranger for the candy. The child lacks capacity—the ability to make a well-reasoned decision in the given circumstances. Second, the damage results not from the child’s choice, nor from his entering the car. The damage results when the adult molests or abducts that child.

One might argue that a sixteen-year-old should know better than to enter a car with a stranger; but should the teenager know better than to follow a caring workplace mentor to another pizzeria? We, as a society, traditionally give children (including adolescents) the benefit of the development doubt until they reach the age of eighteen. They are innocent until adjudged mature. Tennessee law considers Doe incapable of consenting to sex. Why would this same law credit him capable of foretelling his own abuse before it occurs? Second, just as with the younger child, Doe’s damages resulted neither from his requested (and denied) transfer to another location, nor from his consent. They resulted from Abson’s abuse and exploitation of Doe’s immaturity.

Given this discussion of capacity and the distinction between culpability and responsibility, the most likely explanation for the Mama Taori’s holding is that the judge did not credit Doe’s lack of capacity. One without capacity remains faultless, morally innocent and legally shielded. From the perspective of a strict moralist (and perhaps most of society), Doe’s consent destroyed his credibility and negated his subjective offense.²⁶⁰ Because Doe was morally culpable, the judge was willing to let a jury find him legally responsible.

²⁶⁰ See, e.g., Barton v. Bee Line, Inc., 265 N.Y.S. 284, 285 (N.Y. App. Div. 1933) (finding that in 1933 a fifteen-year-old girl could not sue the common carrier for tort damages resulting from a rape when she consented to sex). The Barton court held that the New York statutory rape law served “to protect the virtue of females and to save society from the ills of promiscuous intercourse.” Id. Note that the court did not equate the injury to the women with the injury done to society. With respect to the woman’s injury, the Barton court explained:

It is one thing to say that society will protect itself by punishing those who consort with females under the age of consent; it is another to hold that, knowing the nature of her act, such female shall be rewarded for her indiscretion. Surely public policy—to serve which the statute was adopted—will not be vindicated by recompen sing her for willing participation in that against which the law sought to protect her. The very object of the statute will be frustrated if by a material return for her fall “we should unwarily put it in the power of the female sex to become seducers in their turn.” Smith v. Richards, 29 Conn. 232. Instead of incapacity to consent being a shield to save, it might be a sword to desecrate. The court is of the opinion that a female under the age of eighteen has no cause of action against a male with whom she willingly consorts, if she knows the nature and quality of her act.
Adolescent Sex & the Workplace

4. **Comparative Fault in Mama Taori’s**

The *Mama Taori’s* court revealed its fault-based, moralistic perspective as it dealt with Restatement (Second) of Torts § 892C. The court explained:

As we construe this provision, it eliminates consent as a complete defense to a civil action for damages. It does not, however, prevent the trier-of-fact from considering evidence of consent when it is allocating fault or determining the existence and extent of the plaintiff’s damages.

Deterrence and punishment for illegal acts should be left to the criminal law. The public’s interests are sufficiently protected by the imposition of criminal sanctions. Thus, civil actions for damages should be left to proceed under ordinary tort law principles.261

The court completely sidestepped the provisions of Restatement (Second) of Torts § 892C. If this subsection eliminates consent as a complete defense, because the legislature intended to protect minors from underage sex, how did the court justify removing that shield with respect to damages? The court ignored the shared purposes of both criminal and civil laws. It ignored that criminal law generally protects potential individual victims as well as society. Additionally, the court ignored the deterrent and punitive effects of tort law.

Rather than accept Tennessee’s judgment regarding teen capacity, the court reverted to “ordinary tort law principles” and redirected the litigation into a comparative fault paradigm. The court reasoned: “[c]onsistent with the doctrine of comparative fault, one of these principles is that a mature minor’s conduct, like an adult’s conduct, is relevant with regard to fault and damages.”262

A moralistic perspective also explains the court’s suggestion that Doe’s consent could negate consequential damages. Consent does not disprove damage. It merely releases the tortfeasor from liability, civil legal responsibility. If Abson had offered to punch Doe, and Doe had

---


262. *Id.* Mama Taori’s had responded with a comparative fault defense, directed at the parents as non-parties. The court affirmed the lower court’s decision not to strike that defense as to the parents. *Id.* at *8–*9.
DROBAC_FINAL 5/7/2004  4:22 PM


consented, his consent would not have erased the resulting bloody nose. It would have simply insulated Abson from liability for the broken nose and resulting medical bills. The damage would still exist.

If the Mama Taori’s court had been truly worried about the extent or even existence of damage, it could have directed the litigation in other ways. Defense counsel could have impeached Doe’s testimony regarding his damages by pointing to a lack of corroborating medical or other physical evidence. Defense counsel could have introduced evidence of Doe’s ability to function in other contexts during that same time period. Evidence of persistently good academic evaluations, excellence on an athletic team, and the ability to maintain other friendships might all speak to his good adjustment and the lack of negative impact by the alleged sexual encounter.

If consent has relevance, one might argue that Doe’s consent exacerbated his trauma. The knowledge that he “consented” to homosexual activity and a sexual predator who duped him might enhance his sense of shame and humiliation. However, the introduction of that evidence will inevitably raise prejudices regarding sexuality and morality: good children do not engage in sex. We return again to Hernandez morality: children who have sex do not deserve protection or compensation for their injuries.263

5. Conclusions for Mama Taori’s

The ultimate issue addressed in the interlocutory appeal opinion by the Mama Taori’s court was whether the defense could use Doe’s consent as a legal defense to the FEPS and tort claims. Instead of protecting Doe, the Mama Taori’s court arranged for the trial of Doe. With the harassment trial, he risked what rape victims often

263. Elizabeth Scott and Laurence Steinberg argued for formal recognition of developmental characteristics to avoid racist and other discriminatory results in criminal adjudications. Scott & Steinberg, supra note 257, at 837. They reasoned:

A developmentally-informed boundary constraining decisionmakers represents a collective pre-commitment to recognizing the mitigating character of youth in assigning blame. Otherwise, immaturity often may be ignored when the exigencies of a particular case engender a punitive response . . . . This concern is critical, given the evidence that illegitimate racial and ethnic biases influence attitudes about the punishment of young offenders and that decisionmakers appear to discount the mitigating impact of immaturity in minority youths.

Id. One can see the parallel concern for how sexist attitudes might cause the discounting of immaturity in a sexual harassment case. Decisionmakers might blame the victim without understanding the nature of the target’s developing capacity.

522
Adolescent Sex & the Workplace

experience—a trial of his character and the accusation, “He asked for it!”

Susan Estrich, who argued against the unwelcomeness requirement in sexual harassment Title VII cases, commented upon the relevance of adult consent (or failure to resist adequately):

The [rape] consent standard—and the corresponding inquiries into what a woman did or said, how she “led the man on,” or how she failed adequately to signal her nonconsent—have at least until recently, made successful prosecution of acquaintance rape all but impossible. Where the relationship is “appropriate,” at least to the court’s eyes, judges tend to see sex, not rape. Similarly, in Title VII cases they see sex, not sexual harassment. In both types of cases, they are often wrong. That a certain relationship might be appropriate does not necessarily mean that the man’s behavior has been.

The strongest justification for the welcomeness doctrine is that the rule ensures that consensual workplace sex does not provide the basis for a civil action. The more radical response to this argument is that there is no such thing as truly “welcome” sex between a male boss and a female employee who needs her job. And if there is, then the women who welcome it will not be bringing lawsuits in any event.264

Estrich’s comments prove even more poignant with respect to “consenting” teens. One could argue (and the Tennessee legislature, among others, has decided) that there is no such thing as truly welcome sex between adults and minors below the age of consent. Employment only worsens the calculus. If a mature (but subordinated) woman who needs her job has difficulty refusing a man, think of the trouble an immature teen will experience in trying to refuse the same man.

Estrich concluded that the unwelcomeness requirement serves, in part, to discourage women from filing suit. She noted the empirical studies concerning the prevalence of workplace sexual harassment and pointed to the dearth of lawsuits.265 We began this section on FEPS by noting that only one case documents a complaint by a “consenting” teen.266 Where are the other teens who surely were manipulated? Or those who, unknowingly, said “Yes”?

264. Estrich, supra note 96, at 831.
265. Id. at 833.
266. See supra note 213 and accompanying text.
The *Mama Taori’s* court’s justifications for considering Doe’s consent cannot withstand scrutiny. What the court referred to as the “mature minor” rule is in many cases a misnomer. The “rule of sevens” pretends to mark accurately the development of adolescent maturity. Actually, it is an archaic bright-line rule designed to avoid the onerous and inexact task of evaluating the maturity of a minor at a given point in time that has long since passed.267 The court sidestepped the Restatement’s guidance in § 892C and ignored underlying public policy reasons that justify both criminal and civil law.268 To add insult to injury, the court charged Doe and his parents with the costs of the appeal.269

The *Mama Taori’s* decision could have a chilling effect on future sexual harassment cases. It also demonstrates that a court, struggling with a new issue, might presume Sara capable of consenting and, thereby, prevent a successful sexual harassment suit under Title VII, state FEPS, or relevant tort law.

**B. Tort Cases**

Of the seventeen tort cases, sixteen split evenly on whether they held consent relevant to the tort claims.270 One found that fraud invalidated a minor’s consent, but the court’s reasoning suggests that it might have

267. Scott and Steinberg noted how prejudices threaten any maturity evaluation. They emphasized:

[W]e currently lack the diagnostic tools to evaluate psycho-social maturity reliably on an individualized basis or to distinguish young career criminals from ordinary adolescents who, as adults, will repudiate their reckless experimentation. Litigating maturity on a case-by-case basis is likely to be an error-prone undertaking, with the outcomes determined by factors other than immaturity.  
Scott & Steinberg, supra note 257, at 836–37.

268. See supra Part III.


Adolescent Sex & the Workplace

credited the consent had the fraud not occurred. Several of the cases that ruled consent relevant served as guiding precedent in *Mama Taori’s* and used now familiar reasoning. Those cases that determined consent irrelevant employed reasoning consistent with each other, most adopting the home state’s criminal law assessment of adolescent “consent.”

1. Consent as Irrelevant

In *Doe by Doe v. Greenville Hospital System*, the court ruled that a candy striper, under sixteen-years-old and working in a hospital, could not legally consent to sexual intercourse with a thirty-one-year-old hospital employee. The hospital argued, as had the defense in *Mama Taori’s*, that a criminal statute had no relevance to a battery claim. The hospital further asserted that battery required a nonconsensual touching. The court disagreed, finding that § 16-3-655(3) of the South Carolina Code, which invalidated consent as a defense to a sexual battery against a minor, applied in both criminal and civil contexts. The court held:

As a matter of public policy, the General Assembly has determined a minor under the age of sixteen is not capable of voluntarily consenting to a sexual battery committed by an older

271. *See* Hackett v. Fulton County Sch. Dist., 238 F. Supp. 2d 1330, 1369 (N.D. Ga. 2002) (finding that the lies the science teacher told his male student induced the student’s consent to inappropriate sexual touching). The court held, “[C]onsent to the act by the person affected negates the contact as an actionable tort. ‘As a general rule, there can be no tort committed against a person consenting thereto, if that consent is free and not obtained by fraud, and is the action of a sound mind.’” *Id.* (quoting Mims v. Boland, 138 S.E.2d 902, 906 (1964)).

272. 448 S.E.2d 564 (S.C. Ct. App. 1994). The lower court returned a verdict for Mary in the sum of $545,000. The trial judge reduced the award to $250,000 under the South Carolina Tort Claims Act. *Id.* at 565. The jury rejected a claim by Mary’s father “for loss of custody, companionship, and service.” *Id.* The father’s claim appears to be akin to a seduction claim already discussed. *See supra* note 125.

273. The Workers’ Compensation Act did not preclude Mary Doe’s tort claims because she received only classroom credit and job skills training in exchange for her services. *Greenville Hosp.,* 448 S.E.2d at 567–68. Thus, the court did not consider her an employee. *Id.* Non-employee status also would preclude someone like Mary Doe, engaged in volunteer work, from suing under an applicable state FEPS. *See, e.g.*, O’Connor v. Davis, 126 F.3d 112, 114–16 (2d Cir. 1997) (finding that volunteer student intern did not qualify as an employee under Title VII), *cert. denied*, 522 U.S. 1114 (1998); Lippold v. Duggal Color Projects, Inc., No. 96 CIV 5869(JSM), 1998 WL 13854, at *2 (S.D.N.Y. Jan. 15, 1998) (confirming that unpaid volunteer cannot sue under Title VII because she is not an employee).

Additionally, the court found that the criminal law applied to Mary Doe’s negligent hiring and supervision claims against the hospital.276

The question understandably left unanswered by the Greenville Hospital decision is what the court would have done if Mary Doe had been sixteen, as Sara was. The Greenville Hospital court consistently applied the criminal law presumptions in a civil case, but this application is hardly shocking since the South Carolina Code defined the age of consent at sixteen. Mary Doe was under sixteen, so the criminal presumptions still applied in her case. The law might have forced the dismissal of Mary Doe’s claims had she been only a few months older.277 Again, we notice the relevance of the state-defined age of consent, which may or may not coincide with the age of majority, typically eighteen.

Other courts similarly found that the age of consent, as defined under criminal law, establishes a compelling public policy that under-aged youth lack the ability to consent to sexual conduct.278 In Bostic v. The Smyrna School District,279 the court noted the public policy concerns and reasoned:

It would be a bizarre rule indeed that, for purposes of civil liability, would call a teenager’s “consent” sufficient to make a relationship “welcome” and thus not a basis for civil liability, when the very same relationship is rape under the exacting standards for criminal liability. Bostic’s sadly misguided participation in the affair is no shield from liability for the defendants.280

275. Greenville Hosp., 448 S.E.2d at 566.
276. Id.
277. See, e.g., Teti v. Huron Ins. Co., 914 F. Supp. 1132, 1139–40 (E.D. Pa. 1996) (finding that because the criminal law permits a consent defense when the victim is sixteen, the civil law must recognize the capacity of minors sixteen and older).
278. Wilson v. Tobiassen, 777 P.2d 1379, 1384 (Or. Ct. App. 1989); see also Angie M. v. Hiemstra, 44 Cal. Rptr. 2d 197, 202 (Cal. Ct. App. 1995) (acknowledging “the strong public policy that underlies the Legislature’s enactment of the multiple statutes directed at protecting minors from sexual exploitation”). The Angie M. court distinguished another sexual battery statute, California Civil Code § 1708.5, which required a lack of consent. Angie M., 44 Cal. Rptr. at 202–03. Despite the Angie M. court’s focus on public policy regarding minors, that reference suggests that the court would have denied a claim by Angie M. under § 1708.5.
280. Id. at *6 (citing Mary M. v. N. Lawrence Cmty. Sch. Corp., 131 F.3d 1220, 1227 (11th Cir. 1997), cert. denied, 524 U.S. 952 (1998)). The Bostic court discussed consent in the context of the
Adolescent Sex & the Workplace

In this passage, the Bostic court stressed the lack of logic in an inconsistent treatment of adolescent “consent.” In Robinson v. Moore, the court explained:

The courts in this country have uniformly taken the position that where the defendant’s act constitutes a violation of a statute, such as a rape statute fixing the age of consent to intercourse, which has as its primary purpose the protection of a definite class of persons from their own immaturity of judgment, the plaintiff’s consent is not a defense to a civil action. This language tracks the reasoning of Restatement (Second) of Torts § 892C. Two other courts specifically cited or discussed this provision and the other Restatement sections. Several cases did not make explicit the public policy rationale but held that consent could not constitute a defense under the circumstances alleged.

2. Consent as Relevant

Courts that found consent relevant to the discussion of civil liability based their determinations on several factors. First, as in Mama Taori’s, the courts noted that criminal laws provide no private right of action. As explained earlier, this reason does not clarify why courts should not apply the criminal law presumptions consistently for an existing civil private right of action. Additionally, the absence of a private right of action under a criminal statute fails to address the absence of logic of finding capacity in one system and a lack thereof in another.

Second, courts pointed to the different purposes of the criminal and civil systems. The primary differences relate to the availability of

282. Id. at 583. The court was incorrect that other courts in this country have uniformly adopted this position. See supra Part II.C.3.
damages and the influence of civil liberties afforded older adolescents. With respect to civil liberties, several courts reasoned that because minors engage in certain adult conduct, civil law should assign full legal capacity. In *Cynthia M. v. Rodney E.*[^287] the court listed the abortion right and the right to consent to other types of medical treatment.[^288] The *Cynthia M.* court reasoned:

“Capacity exits when the minor has the ability of the average person to understand and weigh the risks and benefits. Moreover, decisional and statutory law is replete with examples of situations in which a child over the age of fourteen is deemed to have the mental capacity of an adult.”[^289]

Again, this comment confuses a determination of adolescent capacity with other public policy reasons for granting minors the right to engage in these adult activities. As previously noted, just because we permit an adolescent to obtain an abortion does not necessarily mean that we attribute to her the ability of the average adult to weigh the risks and benefits on a regular basis—or even in most circumstances.

The existence and assessment of damages appears to be the major concern of all of these courts. The damage calculus figures into not only the second factor concerning the different purposes of the criminal and civil systems, but also a third factor regarding unfairness. Specifically, courts that permitted consideration of adolescent “consent” emphasized the injustice of granting damages to a participant in a crime and of limiting evidence under those circumstances.

The restriction of evidence concerned the *Mama Taori’s* court and other tribunals. For example, in *Doe by Roe v. Orangeburg County School District*,[^290] the South Carolina Supreme Court allowed admission of plaintiff’s “willing participation” to the sex but only on the issue of damages, not the issue of liability.[^291] The court specifically distinguished the lower court’s holding in *Greenville Hospital*, which found that a child under sixteen did not have the capacity to consent to sex.[^292] In *Orangeburg*, the plaintiff and her parents sued the school district and a teacher who failed to supervise students, during which time a sexual

[^288]: Id. at 97.
[^289]: Id. (quoting PROSSER & KEETON, supra note 148, § 18, at 115) (citation omitted); see also McNamee, 519 S.E.2d at 302 (using the same words as the *Cynthia M.* decision).
[^291]: Id. at 261.
[^292]: Id.
Adolescent Sex & the Workplace

assault occurred. A sixteen-year-old mentally handicapped student allegedly sexually assaulted a fourteen-year-old girl after the coach left them alone in the school gym. The Orangeburg court relied on Barnes v. Barnes, a challenge to the Indiana Rape Shield Statute. Quoting Barnes, the Orangeburg court reasoned:

“Unlike the victim in a criminal case, the plaintiff in a civil damage action is ‘on trial’ in the sense that he or she is an actual party seeking affirmative relief from another party. Such plaintiff is a voluntary participant, with strong financial incentive to shape the evidence that determines the outcome. It is antithetical to principles of fair trial that one party may seek recovery from another based on evidence it selects while precluding opposing relevant evidence on grounds of prejudice.”

This passage highlights the court’s focus on fairness. The court ignored, however, that prejudice regularly justifies the exclusion of probative evidence. Additionally, the court missed the point of exclusion. The main reason for excluding the consent was not the prejudice potentially created, but the minor’s incapacity that rendered the consent legally invalid. This court did not even hesitate to put the consenting minor “on trial.”

The other concern pertaining to fairness centered on the minor’s “willing participation” in the conduct. Just as in Mama Taori’s, many of these cases invoked concepts of comparative fault, contributory negligence, or assumption of risk to deal with the victim’s conduct.

293. Id. at 259.
294. Id.
296. Id. at 1342.
297. Orangeburg, 518 S.E.2d at 261 (quoting Barnes, 603 N.E.2d at 1342).
298. See FED. R. EVID. 403 (directing that “[a]lthough relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice”).
300. See, e.g., Beul v. ASSE Int’l, Inc., 233 F.3d 441, 451 (7th Cir. 2000) (opting for a comparative fault rule); Robinson v. Roberts, 423 S.E.2d 17, 18 (Ga. Ct. App. 1992) (evaluating contributory negligence and assumption of risk to find that a thirteen-year-old can “appreciate dangers of his environment and . . . avoid consequences associated with exposure to such dangers”); LK, 631 So. 2d at 608 (determining that their “analysis must include the principles of comparative fault”).
Many took a moralistic stance, evaluating whether the victim was "innocent" or not. The *Cynthia M.* court explained: “We have emphasized the word ‘innocent’ because we believe there is an important distinction between a party who is injured through no fault of his or her own and an injured party who willingly participated in the offense about which the complaint is made.”

Quoting a 1922 Louisiana Supreme Court decision, the *Cynthia M.* court added, “[T]o recognize the asserted right to recover would be to permit plaintiff to profit by the wrong to which she voluntarily was a party . . . .” Just like the *Mama Taori’s* court, the *Cynthia M.* court ignored the fact that a minor lacks the capacity to consent in the criminal context. The court missed the meaning of the statutory rape charge. The notion that a minor “profits” when she collects money for medical bills associated with a pregnancy, for psychotherapy, or for emotional and physical distress deserves no comment. *Cynthia M.* stands as another classic example in which we blame the victim, this time, a sixteen-year-old.

In *LK v. Reed*, a thirteen-year-old special education student took the blame, or at least a pro rata share of it. A.K., through her estate administrators, sued another student and the school board after A.K. allegedly agreed to engage in sex with an eighteen-year-old special education high school junior. The *LK* court noted:

> [T]he anomaly created by the trial court’s holding . . . necessarily entitles any carnal knowledge victim to civil damages. Under the trial court’s holding, a girl could

---

301. See, e.g., *Cynthia M.*, 279 Cal. Rptr. at 98; *LK*, 631 So. 2d at 607; *Orangeburg*, 518 S.E.2d at 259–60 (noting that the “District proffered testimony tending to dispute the claim Doe was a sweet, innocent young girl with testimony that she had been overheard making sexually explicit statements”).

302. *Cynthia M.*, 279 Cal. Rptr. at 98.

303. *Id.* (quoting Overhultz v. Row, 92 So. 716, 717 (La. 1922)). Having stated that *Overhultz* was “directly on point,” the *Cynthia M.* court then applied reasoning formulated for an adult. *Id.* In its final footnote, the *Cynthia M.* court admitted that the *Overhultz* plaintiff was not a minor. *Id.* at 98 n.14. The court then stated, “However, we are not inclined to dwell on outdated legal fictions concerning the ability of underage females to consent to sex.” *Id.* Commenting upon the current prevalence of underage sexual activity and the problem of pregnancy among unwed teenagers, the court added, “To cling to vestiges of a bygone era, is to ignore the contemporary realities of nature.” *Id.* In this footnote, the *Cynthia M.* court arguably blamed teenagers for their promiscuity and particularly teenaged girls for their non-marital pregnancies. The court’s treatment of their consent and effective denial of their damages may not prove to be the best way to handle these social ills.


305. *Id.* at 607.

306. *Id.* at 605.
Adolescent Sex & the Workplace

provoke a criminal prosecution against a sexual partner and recover damages from him, both as a result of her willful and voluntary actions in consenting to, or instigating, a sexual liaison.307

This passage conjures the specter of the young seductress, luring men to their financial demise. It bears no relation to the reality of a thirteen-year-old special education student with an IQ of between sixty-four and seventy-four.308 The LK court also neglected to consider that any potential sexual partner of such a Lolita309 remained free to reject her advances and spare himself criminal and potential civil liability. This notion of the child harlot, ready to entrap an unsuspecting partner, exemplifies the most dated, sexist notions of women (and girls) as avaricious temptresses.310

307. Id. at 607; see also Orangeburg, 518 S.E.2d at 261 (reasoning that “[t]o prohibit such evidence [of consent] would effectually allow a victim to come in and tell a one-sided version of events, without being subject to any real cross-examination or impeachment as to the damages actually suffered”).

308. LK, 631 So. 2d at 605. Despite its ultimate determination, the court recited other disturbing facts:

[A]t the time of these events A.K. was a 13-year-old girl with minimal intellectual and social skills. She was shy and obedient and had never had a boyfriend. She had a history of seizures for which she took daily medication. Her family was poor in financial assets but rich in religious beliefs. In the year preceding these events, A.K.’s father was involved in an accident which rendered him a paraplegic, and A.K.’s mother donated a kidney to A.K.’s younger sister, a surgery requiring extended visits to New Orleans.

A.K.’s family stress coupled with her age, intellect, and social skills, render her consent, from a legal standpoint, almost meaningless. Accordingly, we assess A.K.’s fault at 5% and reduce the damages awarded to her by that percentage.

Id. at 608.


310. See Marybeth Hamilton Arnold, “The Life of a Citizen in the Hands of a Woman”: Sexual Assault in New York City, 1790–1820, in PASSION AND POWER: SEXUALITY IN HISTORY 35, 40–45 (Kathy Peiss & Christina Simmons eds., 1989) (discussing sexual assault of women at the turn of the nineteenth century and highlighting the popular myth of sexually voracious working class women). Arnold noted the rape of a thirteen-year-old girl who was likened to a harlot. Id. at 42.

Counsel for the defense argued:

[If] anything of an improper nature passed between them, I am inclined to believe that it has been with her consent. The passions may be as warm in a girl of her age as in one of more advanced years, and with very little enticement she may have consented to become his mistress . . . . [It] is said her youth renders it impossible she should have been a lewd girl. Who is acquainted with the dissolute morals of our city, and does not know that females are to be found living in a state of open prostitution at the early ages of 12 and 13 years?

Id. (citing Report of the Trial of Richard D. Croucher, on an Indictment for a Rape of Margaret Miller, on Tuesday, the 8th day of July, 1800, at 15, 18 (New York: 1800)); see also Estelle B. Freedman, “Uncontrolled Desires”: The Response to the Sexual Psychopath, 1920–1960, in PASSION AND POWER, supra, at 199, 212 (explaining that victims of sexual predators were described
In *McNamee v. A.J.W.*, the court suggested that the admissibility of the consent might hinge upon whether the sexual partner was also a minor. This suggestion again reflects an emphasis on comparative fault. As discussed earlier, if neither party possesses the capacity to consent, why should we blame (in a comparative fault scheme) either? Such a policy makes no logical sense. While comparative fault seems inappropriate when applied to minors who lack capacity, the concept hints at a parallel concern—comparative power. The *McNamee* court was perhaps correct (but for the wrong reason) to emphasize the difference between teen-teen and teen-adult consensual sex.

3. *From Comparative Fault to Comparative Power*

Many of the tort cases that found adolescent “consent” relevant to a civil claim for damages favored comparative fault schemes. In contrast, those cases that found consent irrelevant emphasized the power disparity between teenagers and adult partners. Courts attributed enhanced power to factors such as older age and maturity, a position of authority, and a position of confidence or trust. For example, in *Angie M. v. Hiemstra*, the minor consented to sex with a forty-eight-year-old physician with whom she worked. The court found that he “took advantage of his position of authority and of Angie’s confidence in him to cause her to develop a dependent relationship on him ‘in much the manner of the phenomenon of “transference” between a patient and his or her psychotherapist.”

In *Bohrer v. DeHart*, the court rejected a comparative fault instruction after a minister allegedly sexually abused a minor parishioner. The court determined that “dependence, transference and the resulting vulnerability do not cease merely because a child physically...”

---

But see Kathy Peiss, “Charity Girls” and City Pleasures: Historical Notes on Working Class Sexuality, 1880–1920, in *PASSION AND POWER*, supra, at 57, 64 (discussing “charity girls,” working women who traded sex for gifts and attention).

312. Id. at 302–03.
313. Id.
314. See supra Part IV.A.4.
317. Id. at 200.
Adolescent Sex & the Workplace

matures while sexual abuse in secrecy by an adult in a position of trust continues unabated.”319 The Bohrer court determined that consent was inadmissible as a defense because of a power imbalance caused by the minor’s sexual encounters with a religious counselor.320

Other teen-adult relationships, in addition to those involving doctors or ministers, resulted in power disparities acknowledged by courts. A teacher holds a position of authority that permits influence over and creates a power imbalance with an adolescent.321 The Bostic court explained that “Smith’s affair with Bostic cannot be viewed as consensual, given the minority of the student and the relationship of trust and authority which the coach held over her.”322 Additionally, in Orangeburg, a case that found consent relevant to the issue of damages, the court referenced a South Carolina Code criminal provision that prohibited sexual conduct between a minor and an “actor [who] is in a position of familial, custodial, or official authority to coerce the victim to submit or is older than the victim.”323 Thus, both tort and criminal law recognize that a power imbalance creates a greater potential for influence and abuse of a minor.

C. Title IX and § 1983 Cases

Title IX and § 1983 cases raise some of the same issues as the more traditional tort claims but borrow heavily from Title VII jurisprudence. In Benefield v. Board of Trustees of the University of Alabama,324 the court considered whether a fifteen-year-old college student could sue her university under Title IX for sexual harassment by classmates.325 The court distinguished college students from elementary and secondary

319. Id. at 1227.
320. Id. (citing E. Cruz, When the Shepherd Preys on the Flock: Clergy Sexual Exploitation and the Search for Solutions, 19 FLA. ST. U. L. REV. 499 (1991)).
323. Doe by Roe v. Orangeburg County Sch., 518 S.E.2d 259, 260 (S.C. 1999) (reviewing South Carolina Code § 16-3-655(3)).
325. Id. at 1215.
school students. The Benefield court ultimately ruled that “to constitute sexual harassment, the behavior in question must be unwelcome.”

In Mary M. v. North Lawrence Community School Corp., the court took a different position. The Mary M. court reviewed a Title IX claim by an eighth grader against a twenty-one-year-old cafeteria worker. The district court had permitted an instruction regarding whether the plaintiff found the conduct unwelcome. The instruction read:

In order to find in favor of the Plaintiff, you must find first that the alleged sexual advances and/or abuses occurred, and if it did, that the advances and/or abuses were unwelcome by her. Conduct is unwelcome if Diane M. did not solicit or incite it, and if she regarded the conduct as undesirable or offensive. In determining whether the conduct was unwelcome, you should consider such things as Diane M.’s receptiveness to the alleged sexual advances and/or abuse in light of her words, acts and demeanor; her emotional predisposition, if any; the age disparity between her and Andrew Fields; any power disparity between them due to Diane M.’s status as a student and Andrew Field’s status as a school employee.

This instruction tracks the unwelcomeness requirement established for Title VII cases. It also acknowledges the potential power disparity between an adult school worker and a minor.

The appellate court rejected the instruction and held:

While welcomeness is properly a question of fact in the context of Title VII employment discrimination cases, we decline to extend the inquiry to Title IX cases when elementary students are involved. It goes without saying that sexual harassment in

326. Id. at 1218.
327. Id. at 1220. Black’s Law Dictionary defines in loco parentis to mean “[a]cting as a temporary guardian of a child.” BLACK’S LAW DICTIONARY, supra note 121, at 791. Ironically, when defining the noun, Black’s Dictionary refers to “[s]upervision of a young adult by an administrative body such as a university.” Id.
329. 131 F.3d 1220 (7th Cir. 1997).
330. Id. at 1221.
331. Id. at 1223.
332. Id. at 1224.
Adolescent Sex & the Workplace

the workplace is vastly different from sexual harassment in a school setting.333 Despite the obvious differences between school and work, the court listed six reasons why sexual harassment at school deserves a different analysis.334 First, the court noted the greater ability of teachers and school officials to control behavior in the classroom, suggesting that “students look to their teachers for guidance as well as for protection.”335 Second, the court explained that school harassment leaves a “longer lasting impact on its younger victims and institutionalizes sexual harassment as accepted behavior.”336 Third, the court reasoned that while adults can leave a hostile work environment, children can rarely leave school.337 Fourth, the court emphasized that children need a nondiscriminatory environment in which to maximize their intellectual growth. The court stated: “A sexually abusive environment inhibits, if not prevents, the harassed student from developing her full intellectual potential and receiving the most from the academic program.”338 Fifth, the court admonished that schools act in loco parentis while employers do not.339 Finally, the court concluded that “employees are older and (presumably) know how to say no to unwelcome advances, while children may not even understand that they are being harassed.”340

The third and sixth reasons reveal that the Mary M. court assumed a workplace populated by adult workers. If one reviews the court’s reasoning and substitutes adolescent workers for the adults, the court’s analysis weakens, and the two environments (school and work) appear less distinct. For example, one might argue that sexual harassment at work leaves a lasting impact on young workers just as it does on young students. Additionally, as noted previously, many adolescent workers may not understand they are experiencing sexual harassment.341

333. Id. at 1226 (citation omitted). The court noted that in Indiana, where the harassment occurred, elementary school extends through the eighth grade. Id. at 1225 n.6. The court added: “We decline to opine, however, on whether secondary school students can welcome sexual advances in harassment claims arising under Title IX.” Id.
334. See id. at 1226–27.
335. Id. at 1226 (citing Davis v. Monroe County Bd. of Educ., 74 F.3d 1186, 1193 (11th Cir. 1996)).
336. Id. (citing Davis, 74 F.3d at 1193).
337. Id.
338. Id.
339. Id.
340. Id. at 1227.
341. See supra note 62 and accompanying text.
Arguably, the court should have distinguished two different types of sexual harassment—harassment of children and harassment of adults—rather than two different environments. In other words, the unwelcomeness requirement may be appropriate as applied to adults but not as applied to minors, whether they are at school or at work.

1. **Reliance on Criminal Law**

Like many of the tort cases, the *Mary M.* case also raised the relevance of the criminal law’s definition of the age of consent. The *Mary M.* court concluded:

> If elementary school children cannot be said to consent to sex in a criminal context, they similarly cannot be said to welcome it in a civil context. To find otherwise would be incongruous.

An opposite holding would defeat the purposes of Title IX and make children claiming sexual discrimination under Title IX subject to intense scrutiny . . . . If welcomeness were properly an issue for the jury in cases involving elementary students, the very children bringing the suits would be subject to intense scrutiny regarding their responses to their alleged abusers. Trial transcripts would be replete with insinuations that a child dressed or acted in such a manner as to ask for the very conduct she or he is seeking to redress . . . . We decline to allow the inference that an elementary school student is presumed to have not consented to molestation by a twenty-one year old in a criminal case, but welcomed the same conduct in a civil case.

In this passage, the court acknowledged the inconsistency of a failure to apply a criminal law presumption in a civil case. More importantly, the *Mary M.* court understood that once a child’s consent comes into evidence, the child goes on trial. Unlike the *Orangeburg* court, the *Mary M.* court eschewed the notion of putting a child on trial. The *Mary M.* case originated in Indiana. *Mary M.*, 131 F.3d at 1220.
Adolescent Sex & the Workplace

court “refused to transfer the onus on the child to prove that in fact she or he did not welcome the complained-of advances.”

2. An Abuse of Power or Casual Sex?

The Doe v. Taylor Independent School District case again explored the power disparity first acknowledged in the tort cases. It also supports Professor Estrich’s view that how a person characterizes the appropriateness of a relationship will influence whether he calls sexual conduct sex or sexual harassment. The Taylor majority held that a child holds a constitutional right to bodily integrity under the Due Process Clause and permitted a § 1983 claim. While the majority did not specifically address Doe’s consent, the concurrence and dissents raised interesting perspectives.

In his concurrence, Justice Higginbotham described the attitudes of the majority and dissent:

The majority and dissents divide today over the “law,” but that division rests largely on perceptions of the human condition. We have all looked at the same set of facts and come away with quite different perceptions of what transpired between teacher and pupil. The majority sees an exploitation of power and the dissents see casual sex. Make no mistake about it. This case is not about a high school coach who happened to have an affair with a student. It is about abuse of power.

In this passage, the Justice emphasized the teacher’s abuse of power over his student.

346. Mary M., 131 F.3d at 1227.
347. 15 F.3d 443 (5th Cir. 1994).
348. See id.
349. See supra note 264 and accompanying text; see also Oberman, supra note 70, at 34–35.
350. U.S. CONST. amend. XIV. The Taylor court explained:

This circuit held as early as 1981 that “the right to be free of state-occasioned damage to a person’s bodily integrity is protected by the fourteenth amendment guarantee of due process.”

If the Constitution protects a schoolchild against being tied to a chair or against arbitrary paddlings, then surely the Constitution protects a schoolchild from physical sexual abuse—here, sexually fondling a 15-year old school girl and statutory rape—by a public school teacher.

Taylor, 15 F.3d at 450–51 (citations omitted).
351. Taylor, 15 F.3d at 457.
352. Id. at 459 (Higginbotham, J., concurring).
353. See also id. at 460 (Higginbotham, J., concurring) (interpreting “Coach Stroud’s use of his position of authority to pressure and manipulate Doe into sex” as “arbitrary and capricious”).
In contrast, the dissenters focused, in part, on Doe’s consent. Justice Garwood suggested that he would permit the § 1983 claim only if the consenting child was immature. He argued:

It is not clearly established that age fifteen is, per se, sufficiently immature. Plainly Doe was of a sufficient age to bear children. Perhaps that should not be the test and instead arguably a minimum age of sixteen, seventeen, or eighteen would make sense as a bright line for these purposes.

Justice Garwood’s concern regarding the bright-line demarcation raises an interesting question that another dissenter posed later. Justice Edith Jones noted that not all criminal statutory rape laws set the age of consent at fifteen. She explained that in some states the age is lower. She then asked, “[H]as the majority made a constitutional offense of conduct that in some states is not criminal?” This question highlights just one of the problems that results from the American legal system’s failure to implement a consistent policy regarding adolescent “consent” to sexual conduct.

D. Conclusion for Sara?

Does Sara have a justiciable claim for sexual harassment against the theater owner? From this review of criminal, sexual harassment, and tort law, we see that the answer still depends on where she consented and files suit. The claims she brings will also influence the outcome. She faces a long and humiliating trial, the outcome of which is not clear, if she sues in Tennessee under its FEPS. Even a successful statutory rape prosecution against the perpetrator will not assist her there. She has almost no chance for success under antidiscrimination or tort law in those twenty-four states that set the age of consent at sixteen or lower. That number increases to thirty-five if courts in eleven states reject the

354. Id. at 467–68 (Garwood, J., dissenting).
355. Id.
356. Id. at 479 n.8 (Jones, J., dissenting).
357. Id.
358. Id.
359. These twenty-four states include: (14) Hawaii, Maryland, Mississippi, South Carolina; (16) Alabama, Georgia, Indiana, Kansas, Kentucky, Massachusetts, Montana, Nevada, North Carolina, Ohio, Pennsylvania, Rhode Island, West Virginia, and Wyoming; and those states with special fact requirements that do not match Sara’s: (14) Maine, (16) Connecticut, Michigan, Oklahoma, South Dakota, and Vermont. See App. A, infra pp. 546–73.
Adolescent Sex & the Workplace

special facts of her case.\textsuperscript{360} In those states, she will be treated as an adult and her consent will bar most claims. She also has little chance for success in Wisconsin where the age of consent is eighteen but where, in \textit{Michelle T. by Sumpter v. Crozier},\textsuperscript{361} the court determined that consent would defeat a civil battery claim by a minor.\textsuperscript{362}

In two states, California and Louisiana, where the ages of consent are eighteen and seventeen respectively, the law is as murky as it is in Tennessee. In San Diego, California, the \textit{Angie M.} court found consent irrelevant.\textsuperscript{363} However, in \textit{Cynthia M.}, the same San Diego court found it relevant.\textsuperscript{364} Sara’s chance of success there depends completely upon the types of claims she pleads. In Louisiana, the \textit{LK} court found consent relevant, and the same Third Circuit court found it irrelevant two years later in \textit{Landreneau}.\textsuperscript{365} I cannot predict what might happen there for Sara today. Neither can I predict with certainty what will happen in the remaining eleven states.\textsuperscript{366}

V. A SYNTHESIS FOR A FUTURE APPROACH

If some advantage explained the inconsistency in the treatment of adolescent “consent” by the United States civil and criminal systems, one might argue for the maintenance of the status quo. No such advantage exists, however. The systems have evolved inconsistently and incongruously. The only justifying explanation for affording adolescent “consent” legal significance, seen in the civil case law, centers on the

\textsuperscript{360} Without the special facts in Sara’s case, including the age disparity and Cosio’s managerial position, eleven more states join the list: (14) Iowa; (15) Colorado; (16) Alaska, Arkansas, Delaware, Florida, Minnesota, New Hampshire, New Jersey, Utah, and Washington. See App. A, infra pp. 546–73. In Colorado, the \textit{Bohrer} court found consent irrelevant because of the prohibition regarding the sexual violation of a minor by an adult in a position of trust. Bohrer v. DeHart, 943 P.2d 1220, 1227 (Colo. Ct. App. 1996). It is not clear that a Colorado court would find that Sara’s manager held a position of trust. Additionally, in \textit{Bostic}, the plaintiff was fifteen when her relationship with her coach commenced. Bostic v. Smyrna Sch. Dist., No. 01-0261 KAJ, 2003 WL 723262, at *1 (D. Del. Feb. 24, 2003). The Delaware court found that the coach held a position of trust. Thus, the court found consent irrelevant. \textit{Id.} at *6.

\textsuperscript{361} 495 N.W.2d 327 (Wis. 1993).

\textsuperscript{362} \textit{Id.} at 329.


\textsuperscript{366} The age of consent in the remaining eleven states are: (17) Illinois, Missouri, Nebraska, New Mexico, New York, and Texas; (18) Arizona, Idaho, North Dakota, Oregon, and Virginia. See App. A, infra pp. 546–73.
need in civil cases to evaluate the existence and extent of a plaintiff’s damages. As the discussion of *Mama Taori*’s indicates, however, consent does not measure or even indicate damages. Alternative avenues for the exploration of a plaintiff’s injuries exist. Thus, the critical evaluation of the plaintiff’s injuries cannot justify treatment of adolescent “consent” that is inconsistent with the criminal statutory rape scheme or with conclusions based on scientific evidence regarding adolescent development.\(^{367}\)

Public policy concerning the treatment of adolescent “consent” deserves attention and revision. Scientific evidence and social science studies should inform the revision to secure for minors a standard that properly accounts for their constitutional liberties as well as their developmental abilities. I cannot endorse the maintenance of the status quo. It is simply too illogical and provides too little protection for developing teenagers. Moreover, it perpetuates outdated moral judgments concerning non-marital and homosexual sex, as well as stereotypical attitudes about “good” girls and boys.

A. State Statutory Rape Laws and the Rule of Sevens

State statutory rape laws also fail to provide adequate direction for the treatment of adolescent “consent.” The first problem with reliance on statutory rape laws rests on the fact that “the age of consent” varies from state to state and from statute to statute within states.\(^{368}\) “So what?” states rights advocates might ask. They might argue that individual states should enjoy the right to set the age of consent as local judgment dictates. The appropriate response is that scientific evidence does not suggest that minors in Colorado develop physically, emotionally, and mentally any earlier than do those in California or New York. We are talking about developmental capacity, not about moral judgments concerning sex or local attitudes about “deviant” sexual behavior. It makes no logical sense to permit multiple and inconsistent

---

367. See Drobac, *supra* note 3, at 8–9; see also Scott & Steinberg, *supra* note 257, at 928–30. Scott and Steinberg caution:

> Psycho-social development proceeds more slowly than cognitive development. As a consequence, even when adolescent cognitive capacities approximate those of adults, youthful decisionmaking may still differ due to immature judgment. The psycho-social factors most relevant to differences in judgment include: (a) peer orientation, (b) attitudes toward and perception of risk, (c) temporal perspective, and (d) capacity for self-management.

Scott & Steinberg, *supra* note 257, at 813.

determinations of adolescent capacity absent some accurate maturity test. Moreover, Justice Jones’ footnote in Taylor reminds us that these inconsistent statutory rape laws potentially create a constitutional offense in some states where the conduct is not criminal.369

Second, I doubt state legislatures passed statutory rape laws in reliance on medical research regarding adolescent developmental capabilities. These laws have been on the books for decades. We should invest the resources necessary to complete scientific research regarding adolescent capabilities and then implement an appropriate national standard based upon reliable scientific evidence. If states want to choose an age of consent higher than that indicated by scientific evidence, then let the states’ rights versus individual rights debate ensue.

The same reliance on science addresses the archaic rule of sevens and the mature minors doctrine. Until we know that fourteen-year-olds possess the same ability as adults to make reasoned decisions and judgments in unfamiliar circumstances, or under stress, we should presume them incapable and protect our older children from sexual abuse. If we must adopt a bright-line, because we have no sure measure of individual cognitive, neuro-psycho-social maturity, the age of majority better serves us.

B. Teenagers on Trial

One alternative to the bright-line drawing, accomplished by ages of consent, necessitates the case-by-case evaluation of the plaintiff’s maturity. Presumptions, either for or against capacity, also inevitably devolve into such an analysis since the side disadvantaged by the presumption will try to challenge it. Either way, the evaluation puts the teen on trial.

A trial of the teenager’s maturity raises several serious concerns. First, the evaluation typically will occur months (if not years) after the teenager “consented.” Therefore, no objective test can accurately evaluate the teenager’s maturity as it existed at the time of the “consent.”370 Anyone who has bought shoes for a teenager knows that adolescents mature and grow with astonishing rapidity. A teenager who “consents” in April may demonstrate a very different level of maturity than he or she will in December, or years later. Fairness dictates that we

370. See supra note 267.
not use the level of maturity that he or she exhibits during trial or discovery to judge capacity at the time of alleged injurious events.

Second, the indicators that we use to gauge teenager maturity may subject that teenager to humiliating, prejudicial, and perhaps even unconstitutional scrutiny. For example, the defense may attempt to introduce evidence of prior sexual activity to demonstrate a teenager’s maturity. While federal sexual harassment law and rules of evidence\(^\text{371}\) regulate the use of the plaintiff’s sexual history, exceptions allow introduction of that evidence when the plaintiff places the issue in controversy or for other reasons during civil trials.\(^\text{372}\) Defense counsel may argue that the claim of incapacity places the plaintiff’s maturity, and therefore her sexual history, in controversy. Conceivably, a court could allow the introduction of the plaintiff’s history of sexual abuse or incest survival to demonstrate that the plaintiff was aware of the difference between consensual and offensive or coerced sex. Even the threat of such admission or discovery of such evidence might deter some teenagers, and their parents, from prosecuting civil claims.

Third, juveniles risk that a judge and jurors will blame them (the teenagers) for not being more mentally mature when they look physically mature or for bringing their injuries upon themselves with their immaturity. One might argue that jurors can resist the temptation to blame the victim. Think about how many times, however, you have thought to yourself, “Oh, that is so immature!” If behavior is truly the product of immaturity, there should be no associated “blame.” A person cannot help his or her immaturity. Immaturity is a natural stage of development. We do not condemn the mentally challenged for their failure to comprehend. Similarly, we should not blame the immature for their failure to act maturely. However, we do all the time.

Associated with this argument concerning blame, one might suggest that even though we should not censure adolescents for their immaturity, neither should we protect them from the consequences of their behavior. How else are people to learn if they do not suffer the consequences of their behavior? I offer two responses to this question. First, teenagers do not “consent” to what we would otherwise label sexual harassment in isolation. The controversial conduct always involves another person, a co-worker or supervisor. I will guess that this second person is usually

\(^{371}\) FED. R. EVID. 412.

\(^{372}\) See supra notes 297–98, 307 and accompanying text.
Adolescent Sex & the Workplace

That adult has the power to prevent the harm by refusing to become sexually involved with the minor. But for the willing assistance of the adult, the harm would not occur. Thus, I would hold the adult responsible for the consequential damages.

My second response to the imposition of consequences borrows from common parenting wisdom. When a child reaches for the hot pot and burns himself, we do not refuse to treat the burn so that the child will learn a lesson and refrain from grabbing pots on the stove. The pain of the moment should deter the child in the future. Neither do we consider a bandage and dressing a reward for bad behavior. Finally, we would not hesitate to scream “Stop!” at the child to prevent the injury in the first place. Take this simple example and apply it to teenager-adult workplace sex. When a teenager suffers injury from teenager-adult workplace sex, we should not hesitate to compensate the injury. The teenager is injured through no fault of his own because he did not have capacity to consent. We should not view the compensatory monetary payment to be a reward for bad behavior. The money is not an “award,” even if we use that term in legal parlance. Lastly, we should not hesitate to yell “Stop!” to prevent workplace teen-adult sex that traumatizes and injures adolescent workers.

C. Strict Liability—Law Reform and Legal Regulation

The best way to prevent workplace sexual harassment of teenagers involves the implementation of both regulatory mechanisms and statutory reform. First, sexual harassment of minors by adults must become a strict liability offense for which consent is no defense. Lawmakers should amend Title VII, state FEPS, and tort law to account for adolescent workers, their developmental abilities, and the phenomenon of their sexual exploitation. Research evidence regarding adolescent development and sexual harassment of minors should inform and guide statutory reform.

373. If the sexual partner is another adolescent, I would deny recovery since neither teenager has acquired the capacity to consent. Each would bear his or her own costs, unless the employer knew or should have known of the conduct. In that case, I would hold the employer liable under a negligence standard for failing to prevent and cure sexual harassment. Similarly, if the second adolescent is a supervisor, I would impose liability on the employer because the employer has made an adolescent its supervising agent. In that case, the employer should bear the burden of that decision and pay for the resulting damages caused by its agent.

374. I would also support the amendment to Title IX to address the issues raised in this Article as they pertain to sexual harassment of minors by adults at school.
Strict liability will serve many of the public policy concerns raised in this Article. It will ensure the development by employers of appropriate workplace policies and procedures to prevent and cure the sexual predation of adolescents on the job. Strict liability, and the necessarily related exclusion of consent evidence,\textsuperscript{375} will lead to cost redistribution. It will lift the financial burden of sexual harassment injuries from minors and their parents. Sexual predators responsible for the injuries will bear the costs. Employers who regulate the workplace and reap the rewards of adolescent labor will share in carrying the cost burden. Moreover, employers can purchase employment practice liability insurance to spread the cost burden further. Strict tort and antidiscrimination law liability will redress the social grievance of the subordination of this nation’s working youth. Such laws will attack the treatment of adolescents as fungible.

In addition to law reform, education and regulatory mechanisms could address this problem. For example, some states subsidize the employment of young workers.\textsuperscript{376} State employment departments could mandate that in exchange for tax breaks or other benefits received, employers provide special training for adolescent workers regarding workplace rights and sexual harassment. Additionally, states might require that to obtain a permit to work, adolescent applicants must complete a sexual harassment training seminar. State education departments might also add sexual harassment curricula to their health and business administration classes. These are just a few of the possible regulatory approaches.

A strict liability scheme and regulatory mechanisms are compatible with affording adolescents some measure of autonomy and self-determination. An adolescent might still choose to engage in sex with an adult co-worker, who would still run the risk of civil and criminal liability. In essence, this scheme operates like adolescent “consent” to a contract. The sex “contract” is voidable by the adolescent but not void. The adolescent can retract the consent if she realizes during her minority.

\textsuperscript{375}. I would consider making evidence of consent admissible in any second (or successive) trial for money damages if the minor had successfully sued for similar injuries on a prior occasion. Such a rule should satisfy those skeptics who will argue that adolescent seductresses will win their fortunes from multiple unsuspecting employers. Additionally, once an injured adolescent has received her recovery and “learned the lesson,” she should not need the same protection, funded by an employer, as one who is inexperienced and naïve.

\textsuperscript{376}. I thank Professor Cynthia Baker who educated me on this point.
Adolescent Sex & the Workplace

(or shortly thereafter)\textsuperscript{377} that her adult partner took advantage of her “developing capacity” at the workplace.

One might argue that if law imposed strict liability on employers for the sexual harassment of adolescent workers by supervisors, employers would simply stop hiring adolescents. Teenagers would face even higher unemployment rates. I challenge that criticism. Employers can pay adolescent workers less money for the same work performed by adults. Often employers avoid providing part-time adolescent employees benefits that adult workers receive. It makes good business sense to hire young, seasonal, or part-time workers. I doubt that the cost of sexual harassment judgments and training would exceed the financial advantages employers enjoy by employing adolescents—if those employers implemented age-appropriate policies and proper training. Holding employers strictly liable for the sexual exploitation of their minor workers would not shut down the markets of this nation.

Sara and the other Does of this nation deserve our protection as they transition to adulthood. These minors mature and develop full capacity through work and other life experiences. As they mature, we should continue to shield them from the humiliating, devastating trauma of sexual exploitation in the workplace. American tort and sexual harassment laws fail our nation’s working youth. If we ignore the conflict of laws that deny our adolescent children protection and recovery, then their seduction is our sin.

\textsuperscript{377} I would advocate an appropriate limitations period for suit and recovery. \textit{See} Oberman, \textit{supra} note 83, at 782–83 (noting the tolling of the reporting time limitation under some statutory rape statutes until the victim reaches her majority).
APPENDIX A

<table>
<thead>
<tr>
<th>State Code &amp; Comment</th>
<th>Age of Consent</th>
<th>Gender neutral, but must be with member of opposite sex. Perpetrator (Perp): 16 or older; Target (Targ): under 12.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALA. CODE § 13A-6-61</td>
<td>Rape, 1st degree</td>
<td><strong>Phipps (16)</strong></td>
</tr>
<tr>
<td>ALA. CODE § 13A-6-62</td>
<td>Rape, 2nd degree</td>
<td><strong>Phipps (16)</strong></td>
</tr>
<tr>
<td>ALA. CODE § 13A-6-63</td>
<td>Sodomy, 2nd degree</td>
<td></td>
</tr>
<tr>
<td>ALA. CODE § 13A-6-64</td>
<td>Sodomy, 2nd degree</td>
<td></td>
</tr>
<tr>
<td>ALA. CODE § 13A-6-65</td>
<td>Sexual misconduct</td>
<td></td>
</tr>
</tbody>
</table>

I thank Sandie McCarthy-Brown for her excellent research and work on this table. The statutes to which Phipps cites are noted in the comment column with the comment "**Phipps." The age of consent Phipps claims for each state is in parenthesis after his name. Phipps, supra note 75, at 441 app. A. All statutes are current on Westlaw as of February, 2004.
Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
<th>ALASKA STAT. § 11.41.436 Sexual Abuse of a minor, 2nd degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska 16/18</td>
<td></td>
<td>Gender neutral. Perp: 16 or older; Targ: under 13; Targ: 13–15 plus 3 yr. age diff. Perp: 18 or older; Targ: under 18 plus family relationship. Perp: 18 or older plus in a position of authority over target; Targ: under 16. **Phipps (16)</td>
</tr>
</tbody>
</table>

| ALASKA STAT. § 11.41.438 Sexual Abuse of a minor, 3rd degree |
|----------------------|---------------------------------------------------------------------|
| ARIZ. REV. STAT. § 13-1404 Sexual abuse |
| ARIZ. REV. STAT. § 13-1405 Sexual conduct with a minor |
| Arizona 18           | Gender neutral. Targ: 15 or older w/out consent or under 15 if only female breast involved. **Phipps (18) |

**Phipps (16)
### STATE CODE & COMMENT
#### AGE OF CONSENT

<table>
<thead>
<tr>
<th>State</th>
<th>Code &amp; Section</th>
<th>Age of Consent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arkansas</td>
<td>ARK. CODE ANN. § 5-14-103 Rape</td>
<td>Gender neutral. Targ: under 14 plus 3 yr. age diff. Targ: under 18 plus family relation plus 3 yr. age diff.</td>
</tr>
<tr>
<td></td>
<td>ARK. CODE ANN. § 5-14-110 Sexual indecency with a child</td>
<td>Gender neutral. Perp: 18 or older; Targ: under 15 plus 3 yr. age diff.</td>
</tr>
<tr>
<td></td>
<td>ARK. CODE ANN. § 5-14-124 Sexual assault, 1st degree</td>
<td>Gender neutral. Targ: under 18 plus perp. is in position of authority.</td>
</tr>
<tr>
<td></td>
<td>ARK. CODE ANN. § 5-14-126 Sexual assault in the 3rd degree</td>
<td>Gender neutral. Perp: under 18; Targ: under 14 plus 3 yr. age diff.</td>
</tr>
<tr>
<td></td>
<td>ARK. CODE ANN. § 5-14-127 Sexual assault in the 4th degree</td>
<td>Gender neutral. Perp: 20 or older; Targ: under 16. (This is the statutory rape code.) <strong>Phipps (16)</strong></td>
</tr>
<tr>
<td>California</td>
<td>CAL. PENAL CODE § 261.5 Unlawful sexual intercourse with person under 18</td>
<td>Gender neutral. Variations in level of punishment according to age of target. Targ: under 18. <strong>Phipps (18)</strong></td>
</tr>
<tr>
<td>Colorado</td>
<td>COLO. REV. STAT. § 18-3-402 Sexual assault</td>
<td>Gender neutral. Targ: under 15 plus 4 yr. age diff.; Targ: 15–16 plus 10 yr. age diff. <strong>Phipps (17)</strong></td>
</tr>
<tr>
<td></td>
<td>COLO. REV. STAT. § 18-3-404 Unlawful sexual contact</td>
<td>Gender neutral. Targ: under 18.</td>
</tr>
</tbody>
</table>
Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>COLO. REV. STAT.</strong></td>
<td>Gender neutral. Targ: under 15 plus 4 yr. age diff.</td>
</tr>
<tr>
<td>§ 18-3-405</td>
<td>Sexual assault on a child</td>
</tr>
<tr>
<td>§ 18-3-405.3</td>
<td>Sexual assault of a child by one in a position of trust</td>
</tr>
</tbody>
</table>

Connecticut 16/18

<table>
<thead>
<tr>
<th><strong>CONN. GEN. STAT.</strong></th>
<th>Gender neutral. Targ: 13–15 plus 2 yr. age diff.; Targ: under 18 plus perp. is target’s guardian. <strong>Phipps (16)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>§ 53a-71</td>
<td>Sexual assault in the 2nd degree</td>
</tr>
<tr>
<td>§ 53a-70</td>
<td>Sexual assault in the 1st degree</td>
</tr>
<tr>
<td>§ 53a-73a</td>
<td>Sexual assault in the 4th degree</td>
</tr>
</tbody>
</table>

Delaware 16/18

<table>
<thead>
<tr>
<th><strong>DEL. CODE ANN. tit. 11, § 223</strong></th>
<th>Masculine includes the feminine.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DEL. CODE ANN. tit. 11, § 770</strong></td>
<td>Gender neutral. Targ: under 16. Perp: 30 and over; Targ: under 18; Targ: 16–17 plus perp. in position of authority.</td>
</tr>
<tr>
<td>Gender/ Number</td>
<td></td>
</tr>
<tr>
<td>Code Section</td>
<td>Description</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DEL. CODE ANN. tit. 11, § 771</td>
<td>Rape in the 3rd degree</td>
</tr>
<tr>
<td>DEL. CODE ANN. tit. 11, § 773</td>
<td>Rape, 1st degree</td>
</tr>
<tr>
<td>DEL. CODE ANN. tit. 11, § 772</td>
<td>Rape, 2nd degree</td>
</tr>
<tr>
<td>DEL. CODE ANN. tit. 11, § 768</td>
<td>Unlawful sexual contact, 2nd degree</td>
</tr>
<tr>
<td>DEL. CODE ANN. tit. 11, § 762(b)</td>
<td>Gender</td>
</tr>
</tbody>
</table>

| Gender neutral. Targ: under 16 plus 10 yr. age diff. or perp. causes physical injury or serious mental or emotional injury. Perp: 19 or older; Targ: under 14. **Phipps (16) |
| Gender neutral. Perp: 18 older; Targ: under 12; Targ: under 16 plus perp. in position of authority or perp. inflicts serious physical injury. |
| Gender neutral. Targ: under 16 plus perp. causes serious physical injury or uses deadly weapon or dangerous instrument. Perp: 18 or older; Targ: under 12; Targ: under 16 plus perp. in position of authority. |
| Gender neutral. Targ: under 16. “Unless a contrary meaning is clearly required, the male pronoun shall be deemed to refer to both male and female.” |
Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEL. CODE ANN. tit. 11, § 761 Definitions</td>
<td>“(j) A child who has not yet reached his or her sixteenth birthday is deemed unable to consent to a sexual act with a person more than 4 years older than said child. Children who have not yet reached their twelfth birthday are deemed unable to consent to a sexual act under any circumstances.”</td>
</tr>
<tr>
<td>D.C. CODE ANN. § 22-3001 Definitions</td>
<td>“’Child’ means a person who has not yet attained the age of 16 years.” Targ: under 16.</td>
</tr>
<tr>
<td>D.C. CODE ANN. § 22-3009 2nd degree child sexual abuse</td>
<td>Gender neutral. under 16 plus 4 yr. age diff.</td>
</tr>
<tr>
<td>FLA. STAT. ANN. § 800.04 Lewd or lascivious offenses committed upon or in the presence of persons less than 16 years of age</td>
<td>Gender neutral. Diff. punishment if perp. is over/under 18; Targ: 12–15 or under 16.</td>
</tr>
<tr>
<td>FLA. STAT. ANN. § 794.011 Sexual battery</td>
<td>Gender neutral. Perp: 18 or older; Targ: under 12 plus injury to sexual organs; Targ: age 12 or older without consent.</td>
</tr>
<tr>
<td>STATE CODE &amp; COMMENT</td>
<td>AGE OF CONSENT</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>FLA. STAT. ANN. § 794.05 Unlawful sexual activity with</td>
<td>Gender neutral. Perp: 24 or older; Targ: 16 or 17. **Phipps (18)</td>
</tr>
<tr>
<td>certain minors</td>
<td></td>
</tr>
<tr>
<td>Georgia 16</td>
<td></td>
</tr>
<tr>
<td>GA. CODE ANN. § 16-6-3 Statutory Rape</td>
<td>Gender neutral. Different punishment if perp. is under/over 21; Targ:</td>
</tr>
<tr>
<td></td>
<td>under 16. **Phipps (16)</td>
</tr>
<tr>
<td>Hawaii 14/16</td>
<td></td>
</tr>
<tr>
<td>HAW. REV. STAT. § 707-730 Sexual assault, 1st degree</td>
<td>Gender neutral. Targ: under 14; Targ: 14–15 plus 5 yr. age diff. **Phipps</td>
</tr>
<tr>
<td></td>
<td>(16)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>HAW. REV. STAT. § 707-732 Sexual assault, 3rd degree</td>
<td>Gender neutral. Targ: under 14; Targ: age 14–15 plus 5 yr. age diff. **</td>
</tr>
<tr>
<td></td>
<td>Phipps (16)</td>
</tr>
<tr>
<td>Idaho 18</td>
<td></td>
</tr>
<tr>
<td>IDAHO CODE § 18-1506 Sex abuse of child under age 16</td>
<td>Gender neutral. Perp: 18 or older; Targ: under 16.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>IDAHO CODE § 18-1508 Lewd conduct with a minor</td>
<td>Gender neutral. Targ: under 16.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>IDAHO CODE § 18-6101 Rape</td>
<td>This code section specifies that “rape” can only be committed by a male and</td>
</tr>
<tr>
<td></td>
<td>only a female may be a target. Female: under/over 18. **Phipps (18)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>IDAHO CODE § 18-6108 Male Rape</td>
<td>This code section specifies that “male rape” can only be committed by a male</td>
</tr>
<tr>
<td></td>
<td>and only a male may be a target. No age specified.</td>
</tr>
</tbody>
</table>
## Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illinois 17/18</td>
<td>720 ILL. COMP. STAT. 5/12-13 Criminal Sexual Assault</td>
<td>Gender neutral. Targ: under 18 plus perp. is family member. Perp: 17 yrs or older plus in position of authority; Targ: 13–17.</td>
</tr>
<tr>
<td></td>
<td>720 ILL. COMP. STAT. 5/11-6 Indecent Solicitation Of A Child</td>
<td>Gender neutral. Perp: 17 or older; Targ: under 17.</td>
</tr>
<tr>
<td></td>
<td>720 ILL. COMP. STAT. 5/12-15 Sexual abuse</td>
<td>Gender neutral. Perp: under 17; Targ: 9–16; Targ: 13–16 plus less than 5 yr. age diff.</td>
</tr>
<tr>
<td></td>
<td>STATE CODE &amp; COMMENT AGE OF CONSENT</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Indiana</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>16</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IND. CODE § 35-42-4-3 Child molesting</td>
<td>Gender neutral. Diff. punishment if perp. is under/over 21; Targ: under 14.</td>
<td></td>
</tr>
<tr>
<td>IND. CODE § 35-42-4-5 Vicarious sexual gratification</td>
<td>Gender neutral. Perp: 18 or older; Targ: under 14.</td>
<td></td>
</tr>
<tr>
<td>IND. CODE § 35-42-4-6 Child solicitation</td>
<td>Gender neutral. Perp: 18 or older; Targ: under 14.</td>
<td></td>
</tr>
<tr>
<td>IND. CODE § 35-42-4-9 Sexual misconduct with a minor</td>
<td>Gender neutral. Perp: 18 or older; Targ: 14–15. <strong>Phipps (16)</strong></td>
<td></td>
</tr>
<tr>
<td>IND. CODE § 35-42-4-1 Rape</td>
<td>Must be with “member of the opposite sex.” No age specified.</td>
<td></td>
</tr>
<tr>
<td><strong>Iowa</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>14/18</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOWA CODE § 702.5 Definition of “Child”</td>
<td>Targ: under 14.</td>
<td></td>
</tr>
<tr>
<td>IOWA CODE § 709.3 Sexual abuse, 2nd degree</td>
<td>Gender neutral. Targ: under 12.</td>
<td></td>
</tr>
<tr>
<td>IOWA CODE § 709.4 Sexual abuse, 3rd degree</td>
<td>Gender neutral. Targ: 12–13; Targ: 14–15 plus family relationship, position of authority, or 4 yr. age diff. <strong>Phipps (16)</strong></td>
<td></td>
</tr>
<tr>
<td>IOWA CODE § 709.12 Indecent contact</td>
<td>Gender neutral. Perp: 18 or older or 16–17 plus 5 yr. age diff.; Targ: under 14.</td>
<td></td>
</tr>
<tr>
<td><strong>Kansas</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>16</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KAN. STAT. ANN. § 21-3502 Rape</td>
<td>Gender neutral. Targ: under 14.</td>
<td></td>
</tr>
<tr>
<td>STATE CODE &amp; COMMENT</td>
<td>AGE OF CONSENT</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>Kentucky 16</td>
<td>Explanation of legislative age choices (ages 12, 14, &amp; 16). Age of full consent: 16.</td>
<td></td>
</tr>
<tr>
<td>KY. REV. STAT. ANN. § 510.060 Rape, 3rd degree</td>
<td>Gender neutral. Targ: incapable of consenting when under 16.</td>
<td></td>
</tr>
<tr>
<td>State Code &amp; Comment</td>
<td>Age of Consent</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------</td>
<td></td>
</tr>
</tbody>
</table>
| **KY. REV. STAT. ANN. § 510.070**  
| **KY. REV. STAT. ANN. § 510.080**  
| **KY. REV. STAT. ANN. § 510.090**  
| **KY. REV. STAT. ANN. § 510.110**  
| **KY. REV. STAT. ANN. § 510.120**  
| **KY. REV. STAT. ANN. § 510.130**  
| **KY. REV. STAT. ANN. § 446.020**  
Gender & number | Use of the masculine includes the feminine unless otherwise indicated. |
| **Louisiana 17** | Gender neutral. Targ: under 17 plus 3 yr. age diff. |

LA. REV. STAT. ANN. § 14:81  
Indecent behavior with a juvenile
**State Code & Comment**

<table>
<thead>
<tr>
<th>State</th>
<th>Code Description</th>
<th>Gender</th>
<th>Perp:</th>
<th>Targ:</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA.</td>
<td>molestation of a juvenile</td>
<td>neutral</td>
<td>over 17</td>
<td>under 17 plus 3 yr. age diff. plus use of force, intimidation, or position of authority.</td>
</tr>
<tr>
<td>LA.</td>
<td>felony carnal knowledge of a juvenile</td>
<td>neutral</td>
<td>19 yrs or older</td>
<td>12–16. Perp: 17 or older; Targ: 12–14.</td>
</tr>
<tr>
<td>LA.</td>
<td>misdemeanor carnal knowledge of a juvenile</td>
<td>neutral</td>
<td>17–18</td>
<td>under 15 plus 3 yr. age diff.</td>
</tr>
<tr>
<td>ME.</td>
<td>gross sexual assault</td>
<td>neutral</td>
<td>under 14</td>
<td>18 if perp. has supervisory or disciplinary control.</td>
</tr>
<tr>
<td>MD.</td>
<td>rape 2nd degree</td>
<td>neutral</td>
<td>14–15 plus 5 yr. age diff.</td>
<td>21; Targ: 16–17 plus perp. is school personnel in target’s school **Phipps (16)</td>
</tr>
<tr>
<td><strong>Phipps (17)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14/18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Maine**

<table>
<thead>
<tr>
<th>Maine</th>
<th>gross sexual assault</th>
<th>Gender</th>
<th>Targ:</th>
</tr>
</thead>
<tbody>
<tr>
<td>ME.</td>
<td>under 14</td>
<td>neutral</td>
<td>18 if perp. has supervisory or disciplinary control.</td>
</tr>
<tr>
<td>ME.</td>
<td>14–15 plus 5 yr. age diff.</td>
<td></td>
<td>21; Targ: 16–17 plus perp. is school personnel in target’s school **Phipps (16)</td>
</tr>
<tr>
<td><strong>Phipps (16)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Maryland**

<table>
<thead>
<tr>
<th>Maryland</th>
<th>rape 2nd degree</th>
<th>Gender</th>
<th>Targ:</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD.</td>
<td>under 14</td>
<td>neutral</td>
<td>14 plus 4 yr. age diff.</td>
</tr>
<tr>
<td>STATE CODE &amp; COMMENT</td>
<td>AGE OF CONSENT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD. CODE ANN., CRIMINAL LAW § 3-306 Sexual offense 2nd degree</td>
<td>Gender neutral. Targ: under 14 plus 4 yr. age diff.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD. CODE ANN., CRIMINAL LAW § 3-307 Sexual offense, 3rd degree</td>
<td>Gender neutral. Targ: under 14 plus 4 yr. age diff. Perp: 21 or older; Targ: 14–15 **Phipps (15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD. CODE ANN., CRIMINAL LAW § 3-308 Sexual offense, 4th degree</td>
<td>Gender neutral but must involve a vagina. Targ: 14–15 plus 4 yr. age diff.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Massachusetts 16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MASS. GEN. LAWS ANN. ch. 265, § 22A Rape of a child</td>
<td>Gender neutral. Targ: under 16 plus force or threat used.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MASS. GEN. LAWS ANN. ch. 265, § 23 Rape and Abuse of a child</td>
<td>Gender neutral. Targ: under 16. **Phipps (16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MASS. GEN. LAWS ANN. ch. 265, § 24B Assault of child; intent to rape</td>
<td>Gender neutral. Targ: under 16.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MASS. GEN. LAWS ANN. ch. 272 § 4 Inducing persons under eighteen to have sexual intercourse</td>
<td>Gender neutral. Targ: under 18 but must be of “chaste life.” [I do not consider this a protective statute since it leaves the target open to a trial regarding her sexual history.]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michigan 16/18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MICH. COMP. LAWS § 750.520b Criminal Sexual Conduct, 1st degree</td>
<td>Gender neutral. Targ: under 13; Targ: 13–15 plus perp. in position of authority.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MICH. COMP. LAWS</strong></td>
<td></td>
</tr>
<tr>
<td>§ 750.520c</td>
<td>Gender neutral. Targ: under 13; Targ: 13–15 plus Perp. in position of authority.</td>
</tr>
<tr>
<td>Criminal Sexual</td>
<td></td>
</tr>
<tr>
<td>Conduct, 2nd degree</td>
<td></td>
</tr>
<tr>
<td>§ 750.520d</td>
<td>Gender neutral. Targ: 13–15; Targ: 16–17 plus perp. is school personnel in target’s school. <strong>Phipps (16)</strong></td>
</tr>
<tr>
<td>Criminal Sexual</td>
<td></td>
</tr>
<tr>
<td>Conduct, 3rd degree</td>
<td></td>
</tr>
<tr>
<td>§ 750.520e</td>
<td>Gender neutral. Targ: 13–15 plus 5 yr. age diff.; Targ: 16–17 plus perp. is in position of authority; Targ: under 16 plus perp. has “significant relationship” to target plus force or coercion used.</td>
</tr>
<tr>
<td>Criminal Sexual</td>
<td></td>
</tr>
<tr>
<td>Conduct 4th degree</td>
<td></td>
</tr>
<tr>
<td><strong>MINN. STAT. § 609.342</strong></td>
<td>Gender neutral. Targ: under 13 plus 36 mo. age diff.; Targ: 13–15 plus 48 mo. age diff. plus perp. is in position of authority; Targ: under 16 plus perp. has “significant relationship” to target plus force or coercion used.</td>
</tr>
<tr>
<td>Criminal Sexual</td>
<td></td>
</tr>
<tr>
<td>Conduct, 1st degree</td>
<td></td>
</tr>
<tr>
<td><strong>MINN. STAT. § 609.343</strong></td>
<td>Gender neutral. Targ: under 13 plus 36 mo. age diff.; Targ: 13–15 plus 48 mo. age diff. plus perp. is in position of authority; Targ: under 16 plus perp. has “significant relationship” to target plus force or coercion used.</td>
</tr>
<tr>
<td>Criminal Sexual</td>
<td></td>
</tr>
<tr>
<td>Conduct, 2nd degree</td>
<td></td>
</tr>
<tr>
<td>STATE CODE &amp; COMMENT AGE OF CONSENT</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------</td>
<td></td>
</tr>
</tbody>
</table>
| **MINN. STAT. § 609.344**  
Criminal Sexual Conduct, 3rd degree | Gender neutral. Targ:  
**Phipps (16)** |
| **MINN. STAT. § 609.345**  
Criminal Sexual Conduct, 4th degree | Gender neutral. Targ:  
under 13 plus 36 mo. age diff.; Targ: 13–15 plus 48 mo. age diff. or perp. is in position of authority; Targ: 16–17 plus 48 mo. age diff. plus perp. is in position of authority; Targ: 16–17 plus significant relationship to perp. |
| **MINN. STAT. § 609.3451**  
Criminal Sexual Conduct, 5th degree | Gender neutral. Targ:  
under 16. |
| **MINN. STAT. § 609.352**  
Solicitation of Children to Engage in Sexual Conduct | Gender neutral. Perp: 18 or older; Targ: 15 and under. |
| Mississippi  
14/16 | MISS. CODE ANN.  
§ 97-3-65  
Statutory Rape | Gender neutral. Perp: 17 or older; Targ: 14–15 plus 36 mo. age diff.; Targ: under 14 plus 24 mo. age diff. **Phipps (17)** |
Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Missouri 17</strong></td>
<td></td>
</tr>
<tr>
<td>MO. REV. STAT. § 566.032</td>
<td>Statutory Rape, 1st degree</td>
</tr>
<tr>
<td>MO. REV. STAT. § 566.034</td>
<td>Statutory Rape, 2nd degree</td>
</tr>
<tr>
<td>MO. REV. STAT. § 566.064</td>
<td>Statutory Sodomy, 2nd degree</td>
</tr>
<tr>
<td><strong>Montana 16</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Nebraska 17</strong></td>
<td></td>
</tr>
<tr>
<td>NEB. REV. STAT. § 28-319</td>
<td>Sexual assault, 1st degree</td>
</tr>
<tr>
<td>NEB. REV. STAT. § 28-320</td>
<td>Sexual assault, 2nd or 3rd degree</td>
</tr>
<tr>
<td>NEB. REV. STAT. § 28-320.01</td>
<td>Sexual assault of a child</td>
</tr>
<tr>
<td>NEB. REV. STAT. § 28-805</td>
<td>Debauching a minor</td>
</tr>
<tr>
<td>State</td>
<td>Code and Comment</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Nevada</td>
<td>NEV. REV. STAT. § 200.368 Statutory sexual seduction: Penalties</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Specifies age of perpetrator.</td>
</tr>
<tr>
<td></td>
<td><strong>Phipps (16)</strong></td>
</tr>
<tr>
<td></td>
<td>NEV. REV. STAT. § 200.364 Definitions</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Perp: 18 or older; Targ: under 16.</td>
</tr>
<tr>
<td></td>
<td><strong>Phipps (16)</strong></td>
</tr>
<tr>
<td>New Hampshire</td>
<td>N.H. REV. STAT. ANN. § 632-A:2 Aggravated felonious sexual assault</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Targ: under 13; Targ: 13–15 plus perp. lives in same household or</td>
</tr>
<tr>
<td></td>
<td>family relation; Targ: 13–17 plus perp. in position of authority.</td>
</tr>
<tr>
<td></td>
<td>N.H. REV. STAT. ANN. § 632-A:3 Felonious sexual assault</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Targ: 13–15 yrs plus penetration; Targ: under 13, contact only.</td>
</tr>
<tr>
<td></td>
<td><strong>Phipps (16)</strong></td>
</tr>
<tr>
<td></td>
<td>N.H. REV. STAT. ANN. § 632-A:4 Sexual assault</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Targ: 13 and over; Targ: 13–15 plus age diff. of 3 yrs. or fewer.</td>
</tr>
<tr>
<td>New Jersey</td>
<td>N.J. STAT. ANN. § 2C:14-1 Definitions</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. No age specified.</td>
</tr>
<tr>
<td></td>
<td>N.J. STAT. ANN. § 2C:14-2 Sexual assault</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Targ: under 13; Targ: 13–15 plus family relation or perp. in</td>
</tr>
<tr>
<td></td>
<td>position of authority or 4 yr. age diff.; Targ: 16–17 plus family relation or</td>
</tr>
<tr>
<td></td>
<td>perp. in position of authority.</td>
</tr>
<tr>
<td></td>
<td><strong>Phipps (16)</strong></td>
</tr>
<tr>
<td></td>
<td>N.J. STAT. ANN. § 2C:14-4 Lewdness (exposure)</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Targ: under 13 plus 4 yr. age diff.</td>
</tr>
</tbody>
</table>
## Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>New Mexico 17/18</strong></td>
<td>Gender neutral. Targ: under 13; Targ: 13–18 plus position of authority. Perp: 18 or older plus 4 yr. age diff.; Targ: 13–16; Targ: 13–18 plus perp. is school personnel in target’s school. <strong>Phipps (16)</strong></td>
</tr>
<tr>
<td>N.M. STAT. ANN. § 30-9-11 Criminal sexual penetration</td>
<td>Gender neutral. Targ: 18 and over plus no consent.</td>
</tr>
<tr>
<td>N.M. STAT. ANN. § 30-9-12 Criminal sexual contact</td>
<td>Gender neutral. Targ: under 13; Targ: 13–18 plus perp. in position of authority.</td>
</tr>
<tr>
<td>N.M. STAT. ANN. § 30-9-13 Criminal sexual contact w/ a minor</td>
<td>Gender neutral. Targ: under 17.</td>
</tr>
<tr>
<td><strong>New York 17</strong></td>
<td>Gender neutral. Perp: 21 or older; Targ: under 17. <strong>Phipps (17)</strong></td>
</tr>
<tr>
<td>N.Y. PENAL LAW § 130.05 Sexual offense: lack of consent</td>
<td>Gender neutral. Perp: 18 or older; Targ: under 15.</td>
</tr>
<tr>
<td>N.Y. PENAL LAW § 130.25 Rape, 3rd degree</td>
<td>Gender neutral. Perp: 21 or older; Targ: under 17.</td>
</tr>
<tr>
<td>N.Y. PENAL LAW § 130.30 Rape, 2nd degree</td>
<td>Gender neutral. Perp: 18 or older; Targ: under 15.</td>
</tr>
<tr>
<td>N.Y. PENAL LAW § 130.40 Criminal sex act, 3rd degree</td>
<td>Gender neutral. Perp: 21 or older; Targ: under 17.</td>
</tr>
<tr>
<td>N.Y. PENAL LAW § 130.45 Criminal sex act, 2nd degree</td>
<td>Gender neutral. Perp: 18 or older; Targ: under 15 plus 4 yr. age diff.</td>
</tr>
<tr>
<td>STATE CODE &amp; COMMENT</td>
<td>AGE OF CONSENT</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------</td>
</tr>
</tbody>
</table>
| § 130.50  
Criminal sex act, 1st degree | |
| § 130.70  
Aggravated sexual abuse, 1st degree | Gender neutral. Targ: under 11. (Legislation changing ages is pending: 2003 New York Senate Bill No. 5278) |
| § 130.60  
Sexual abuse, 2nd degree | Gender neutral. Targ: under 14. |
| § 130.65  
Sexual abuse, 3rd degree | Gender neutral. Targ: under 11. (Legislation changing ages is pending: 2003 New York Senate Bill No. 5278) |
| § 130.66  
Aggravated sexual abuse, 3rd degree | Gender neutral. Targ: under 11. (Legislation changing ages is pending: 2003 New York Senate Bill No. 5278) |
| § 130.67  
Aggravated sexual abuse, 2nd degree | Gender neutral. Targ: under 11. (Legislation changing ages is pending: 2003 New York Senate Bill No. 5278) |
| **North Carolina 16** | Gender neutral. Targ: 13–15 plus 6 yr. age diff. **Phipps (16)** |
| § 14-27.7A  
Statutory rape or sexual offense | |
| § 14-27.2  
Rape, 1st degree | Gender neutral. Perp: 12 or older; Targ: under 13 plus 4 yr. age diff. |
### Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N.C. GEN. STAT.</strong></td>
<td>Gender neutral. Perp: 12 or older; Targ: under 13 plus 4 yr. age diff.</td>
</tr>
<tr>
<td>§ 14-27.4</td>
<td>Sexual offense, 1st degree</td>
</tr>
<tr>
<td>North Dakota</td>
<td>Gender neutral. Targ: under 15.</td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
<tr>
<td><strong>N.D. CENT. CODE</strong></td>
<td>Gender neutral. Targ: under 15.</td>
</tr>
<tr>
<td>§ 12.1-20-03</td>
<td>Gross sexual imposition</td>
</tr>
<tr>
<td>§ 12.1-20-03.1</td>
<td>Continuous sexual abuse of child</td>
</tr>
<tr>
<td><strong>N.D. CENT. CODE</strong></td>
<td>Gender neutral. Perp: 18 or older; Targ: 15–17. Perp: 18 or older; Targ: under 15. Perp: 22 or older; Targ: 15–17. <strong>Phipps (15)</strong></td>
</tr>
<tr>
<td>§ 12.1-20-05</td>
<td>Corruption or solicitation of minor</td>
</tr>
<tr>
<td><strong>Ohio</strong></td>
<td>Gender neutral. Perp: 18 or older; Targ: 15–17.</td>
</tr>
<tr>
<td>16</td>
<td>“Minor” means under age 18; “juvenile” means under 18.</td>
</tr>
<tr>
<td>§ 2907.01</td>
<td>Definitions</td>
</tr>
<tr>
<td><strong>OHIO REV. CODE ANN.</strong></td>
<td>Gender neutral. Perp: 18 or older; Targ: 13–15. <strong>Phipps (16)</strong></td>
</tr>
<tr>
<td>§ 2907.04</td>
<td>Unlawful sex. conduct with a minor</td>
</tr>
<tr>
<td><strong>OHIO REV. CODE ANN.</strong></td>
<td>Gender neutral. Perp: 18 or older; Targ: 13–15 plus 4 yr. age diff.</td>
</tr>
<tr>
<td>§ 2907.05</td>
<td>Gross sexual imposition</td>
</tr>
<tr>
<td><strong>OHIO REV. CODE ANN.</strong></td>
<td></td>
</tr>
<tr>
<td>§ 2907.06</td>
<td>Sexual imposition</td>
</tr>
</tbody>
</table>

**Phipps (15)**
<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oklahoma</strong></td>
<td></td>
</tr>
<tr>
<td>16/18</td>
<td></td>
</tr>
<tr>
<td>OKLA. STAT. tit. 21</td>
<td></td>
</tr>
<tr>
<td>§ 1111</td>
<td>Definitions</td>
</tr>
<tr>
<td>Gender neutral rape. Targ: under 16; Targ: 16–18 plus perp. is school personnel in target’s school. <strong>Phipps (15)</strong></td>
<td></td>
</tr>
<tr>
<td>OKLA. STAT. tit. 21</td>
<td></td>
</tr>
<tr>
<td>§ 1114</td>
<td>Rape, 1st degree &amp; 2nd degree</td>
</tr>
<tr>
<td>Gender neutral. Perp: 18 or older; Targ: under 14.</td>
<td></td>
</tr>
<tr>
<td>OKLA. STAT. tit. 21</td>
<td></td>
</tr>
<tr>
<td>§ 1123</td>
<td>Lewd or indecent propositions or acts to a child</td>
</tr>
<tr>
<td>Gender neutral. Targ: under 16.</td>
<td></td>
</tr>
<tr>
<td><strong>Oregon</strong></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
<tr>
<td>OR. REV. STAT.</td>
<td></td>
</tr>
<tr>
<td>§ 163.315</td>
<td>Incapacity to consent</td>
</tr>
<tr>
<td>Gender neutral. Targ: under 18.</td>
<td></td>
</tr>
<tr>
<td>OR. REV. STAT.</td>
<td></td>
</tr>
<tr>
<td>§ 163.355</td>
<td>Rape, 3rd degree</td>
</tr>
<tr>
<td>Gender neutral. Targ: under 16.</td>
<td></td>
</tr>
<tr>
<td>OR. REV. STAT.</td>
<td></td>
</tr>
<tr>
<td>§ 163.365</td>
<td>Rape, 2nd degree</td>
</tr>
<tr>
<td>Gender neutral. Targ: under 14.</td>
<td></td>
</tr>
<tr>
<td>OR. REV. STAT.</td>
<td></td>
</tr>
<tr>
<td>§ 163.375</td>
<td>Rape, 1st degree</td>
</tr>
<tr>
<td>Gender neutral. Targ: under 12; Targ: under 16 plus certain family relationships.</td>
<td></td>
</tr>
<tr>
<td>OR. REV. STAT.</td>
<td></td>
</tr>
<tr>
<td>§ 163.385</td>
<td>Sodomy, 3rd degree</td>
</tr>
<tr>
<td>Gender neutral. Targ: under 16.</td>
<td></td>
</tr>
<tr>
<td>OR. REV. STAT.</td>
<td></td>
</tr>
<tr>
<td>§ 163.395</td>
<td>Sodomy, 2nd degree</td>
</tr>
<tr>
<td>Gender neutral. Targ: under 14.</td>
<td></td>
</tr>
<tr>
<td>OR. REV. STAT.</td>
<td></td>
</tr>
<tr>
<td>§ 163.405</td>
<td>Sodomy, 1st degree</td>
</tr>
<tr>
<td>Gender neutral. Targ: under 12; Targ: under 16 plus certain family relationships.</td>
<td></td>
</tr>
</tbody>
</table>
# Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oregon</strong></td>
<td></td>
</tr>
<tr>
<td>OR. REV. STAT. § 163.411</td>
<td>Unlawful sex. Penetration, 1st degree</td>
</tr>
<tr>
<td>OR. REV. STAT. § 163.415</td>
<td>Sexual abuse, 3rd degree</td>
</tr>
<tr>
<td>OR. REV. STAT. § 163.435</td>
<td>Contributing to the sexual delinquency of a minor</td>
</tr>
<tr>
<td><strong>Pennsylvania</strong></td>
<td></td>
</tr>
<tr>
<td>18 PA. CONS. STAT. § 3121</td>
<td>Rape of a child</td>
</tr>
<tr>
<td>18 PA. CONS. STAT. § 3122.1</td>
<td>Statutory sexual assault</td>
</tr>
<tr>
<td>18 PA. CONS. STAT. § 3125</td>
<td>Aggravated indecent assault</td>
</tr>
<tr>
<td>18 PA. CONS. STAT. § 3126</td>
<td>Indecent assault</td>
</tr>
<tr>
<td>STATE CODE &amp; COMMENT</td>
<td>AGE OF CONSENT</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Rhode Island 16</td>
<td>R.I. GEN. LAWS § 11-37-6 Sexual assault 3rd degree</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Targ: 14–15. **Phipps (16)</td>
</tr>
<tr>
<td></td>
<td>R.I. GEN. LAWS § 11-37-8.1 Child molestation 1st degree</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Targ: 14 and under.</td>
</tr>
<tr>
<td></td>
<td>R.I. GEN. LAWS § 11-37-8.3 Child molestation 2nd degree</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Targ: 14 and under.</td>
</tr>
<tr>
<td>South Carolina 14/16</td>
<td>S.C. CODE ANN. § 16-3-655 Criminal sexual conduct with minors</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Targ: under 11; Targ: 11–14; 14–15 plus perp. is in a position of familial, custodial or official authority. **Phipps (16)</td>
</tr>
<tr>
<td>South Dakota 16/18</td>
<td>S.D. CODIFIED LAWS § 22-22-1 Rape defined</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Targ: under 10; Targ: 10–15 plus 3 yr. age diff.; Targ: 10–17 plus perp. is parent’s spouse or former spouse. **Phipps (16)</td>
</tr>
<tr>
<td></td>
<td>S.D. CODIFIED LAWS § 22-22-7 Sexual contact with a child</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Perp: 16 or older; Targ: under 16.</td>
</tr>
<tr>
<td></td>
<td>S.D. CODIFIED LAWS § 22-22-7.3 Sexual contact with a child; misdemeanor</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Perp: 16 or under; Targ: under 16.</td>
</tr>
<tr>
<td></td>
<td>S.D. CODIFIED LAWS § 22-22-24.5 Solicitation of a minor</td>
</tr>
<tr>
<td></td>
<td>Gender neutral. Perp: 18 or older; Targ: under 18.</td>
</tr>
</tbody>
</table>
## Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TENNESSEE</strong> 13/18</td>
<td></td>
</tr>
<tr>
<td>aggravated sexual battery</td>
<td></td>
</tr>
<tr>
<td>statutory rape</td>
<td><strong>Phipps (18)</strong></td>
</tr>
<tr>
<td>rape of a child</td>
<td></td>
</tr>
<tr>
<td>TENN. CODE. ANN. § 39-13-527</td>
<td>Gender neutral. Targ: age 13–17 plus perp. in supervisory or disciplinary position.</td>
</tr>
<tr>
<td>authority figure, sexual battery</td>
<td></td>
</tr>
<tr>
<td><strong>TEXAS</strong> 17</td>
<td></td>
</tr>
<tr>
<td>sexual assault</td>
<td><strong>Phipps (17)</strong></td>
</tr>
<tr>
<td>aggravated sexual assault</td>
<td></td>
</tr>
<tr>
<td>TEX. PENAL CODE ANN. § 21.01</td>
<td>Gender neutral. No age specified.</td>
</tr>
<tr>
<td>definitions</td>
<td></td>
</tr>
<tr>
<td>indecency with a child</td>
<td></td>
</tr>
<tr>
<td><strong>UTAH</strong> 16/18</td>
<td></td>
</tr>
<tr>
<td>unlawful sex activity with a minor</td>
<td></td>
</tr>
<tr>
<td>sexual abuse of a minor</td>
<td></td>
</tr>
<tr>
<td>STATE CODE &amp; COMMENT</td>
<td>AGE OF CONSENT</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------</td>
</tr>
<tr>
<td><strong>UTAH CODE ANN.</strong></td>
<td></td>
</tr>
<tr>
<td>§ 76-5-401.2</td>
<td>Gender neutral. Targ: 16–17 plus 10 yr. age diff.</td>
</tr>
<tr>
<td>Unlawful sexual conduct with a 16 or 17 old</td>
<td></td>
</tr>
<tr>
<td><strong>UTAH CODE ANN.</strong></td>
<td></td>
</tr>
<tr>
<td>§ 76-5-402</td>
<td>Gender neutral. No age specified. <strong>Phipps (16)</strong></td>
</tr>
<tr>
<td>Rape</td>
<td></td>
</tr>
<tr>
<td><strong>UTAH CODE ANN.</strong></td>
<td></td>
</tr>
<tr>
<td>§ 76-5-402.1</td>
<td>Gender neutral. Targ: under 14.</td>
</tr>
<tr>
<td>Rape of a child</td>
<td></td>
</tr>
<tr>
<td><strong>UTAH CODE ANN.</strong></td>
<td></td>
</tr>
<tr>
<td>§ 76-5-402.3</td>
<td>Gender neutral. Targ: under 14.</td>
</tr>
<tr>
<td>Object rape of a child</td>
<td></td>
</tr>
<tr>
<td><strong>UTAH CODE ANN.</strong></td>
<td></td>
</tr>
<tr>
<td>§ 76-5-403.1</td>
<td>Gender neutral. Targ: under 14.</td>
</tr>
<tr>
<td>Sodomy of a child</td>
<td></td>
</tr>
<tr>
<td><strong>UTAH CODE ANN.</strong></td>
<td></td>
</tr>
<tr>
<td>§ 76-5-404</td>
<td>Gender neutral. Targ: 14 and over.</td>
</tr>
<tr>
<td>Forcible sexual abuse</td>
<td></td>
</tr>
<tr>
<td><strong>UTAH CODE ANN.</strong></td>
<td></td>
</tr>
<tr>
<td>§ 76-5-404.1</td>
<td>Gender neutral. Targ: under 14.</td>
</tr>
<tr>
<td>Sexual abuse of a child, aggravated</td>
<td></td>
</tr>
<tr>
<td><strong>UTAH CODE ANN.</strong></td>
<td></td>
</tr>
<tr>
<td>§ 76-5-406</td>
<td>Gender neutral. Targ: under 14; Targ: under 18 plus certain family relationships; Targ: 14–17 plus 3 yr. age diff.</td>
</tr>
<tr>
<td>Sexual offense without consent</td>
<td></td>
</tr>
<tr>
<td>Vermont 16/18</td>
<td></td>
</tr>
<tr>
<td>VT. STAT. ANN. tit. 13, § 3252</td>
<td>Gender neutral. Targ: under 16; Targ: under 18 plus certain family relationships. <strong>Phipps (16)</strong></td>
</tr>
<tr>
<td>Sexual assault</td>
<td></td>
</tr>
<tr>
<td>Aggravated sexual assault</td>
<td></td>
</tr>
</tbody>
</table>
Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Virginia 18</strong></td>
<td></td>
</tr>
<tr>
<td>VA. CODE ANN.</td>
<td>Gender neutral. Targ:</td>
</tr>
<tr>
<td>§ 18.2-61 Rape</td>
<td>under 13.</td>
</tr>
<tr>
<td>VA. CODE ANN.</td>
<td>Gender neutral. Targ:</td>
</tr>
<tr>
<td>§ 18.2-63 Carnal knowledge of a child</td>
<td>13–14.</td>
</tr>
<tr>
<td>VA. CODE ANN.</td>
<td>Gender neutral. Targ:</td>
</tr>
<tr>
<td>§ 18.2-67.1 Forcible sodomy</td>
<td>under 13.</td>
</tr>
<tr>
<td>VA. CODE ANN.</td>
<td>Gender neutral. Targ:</td>
</tr>
<tr>
<td>§ 18.2-67.2 Object sexual penetration</td>
<td>under 13.</td>
</tr>
<tr>
<td>VA. CODE ANN.</td>
<td>Gender neutral. Targ:</td>
</tr>
<tr>
<td>§ 18.2-67.3 Aggravated sexual battery</td>
<td>under 13. Plus use of force or threat.</td>
</tr>
<tr>
<td>VA. CODE ANN.</td>
<td>Gender neutral. Perp: 18 or older; Targ: “child” 15 or older. **Phipps (18)</td>
</tr>
<tr>
<td>§ 18.2-371 Causing or encouraging acts rendering children delinquent, abused, etc.; penalty; abandoned infant</td>
<td></td>
</tr>
<tr>
<td><strong>Washington 16/18</strong></td>
<td></td>
</tr>
<tr>
<td>WASH. REV. CODE</td>
<td>Gender neutral. Targ:</td>
</tr>
<tr>
<td>§ 9A.44.073 Rape of a child, 1st degree</td>
<td>under 12 plus 24 mo. age diff.</td>
</tr>
<tr>
<td>WASH. REV. CODE</td>
<td>Gender neutral. Targ:</td>
</tr>
<tr>
<td>§ 9A.44.076 Rape of a child, 2nd degree</td>
<td>age 12–13 plus 36 mo. age diff.</td>
</tr>
<tr>
<td>WASH. REV. CODE</td>
<td>Gender neutral. Targ:</td>
</tr>
<tr>
<td>§ 9A.44.079 Rape of a child, 3rd degree</td>
<td>age 14–15 plus 48 mo. age diff. **Phipps (16)</td>
</tr>
<tr>
<td>STATE CODE &amp; COMMENT</td>
<td>AGE OF CONSENT</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------</td>
</tr>
<tr>
<td><strong>WASH. REV. CODE</strong></td>
<td></td>
</tr>
<tr>
<td>§ 9A.44.083</td>
<td>Gender neutral. Targ: under 12 plus 36 mo. age diff.</td>
</tr>
<tr>
<td>Child molestation, 1st degree</td>
<td></td>
</tr>
<tr>
<td>§ 9A.44.086</td>
<td>Gender neutral. Targ: 12–13 plus 36 mo. age diff.</td>
</tr>
<tr>
<td>Child molestation, 2nd degree</td>
<td></td>
</tr>
<tr>
<td>§ 9A.44.089</td>
<td>Gender neutral. Targ: 14–15 plus 48 mo. age diff.</td>
</tr>
<tr>
<td>Child molestation, 3rd degree</td>
<td></td>
</tr>
<tr>
<td>§ 9A.44.093</td>
<td>Gender neutral. Targ: 16–17 plus 60 mo. age diff., perp. in significant relationship with target, and perp. in supervisory position over target.</td>
</tr>
<tr>
<td>Sexual misconduct with a minor, 1st degree</td>
<td></td>
</tr>
<tr>
<td>§ 9A.44.096</td>
<td>Gender neutral. Targ: 16–17.</td>
</tr>
<tr>
<td>Sexual misconduct with a minor, 2nd degree</td>
<td></td>
</tr>
<tr>
<td><strong>W. VA. CODE</strong></td>
<td></td>
</tr>
<tr>
<td>§ 61-8B-2</td>
<td>Gender neutral. Targ: under 16.</td>
</tr>
<tr>
<td>Lack of consent</td>
<td></td>
</tr>
<tr>
<td>§ 61-8B-3</td>
<td>Gender neutral. Perp: 14 or older; Targ: 11 and under.</td>
</tr>
<tr>
<td>Sexual assault, 1st degree</td>
<td></td>
</tr>
<tr>
<td>§ 61-8B-5</td>
<td>Gender neutral. Perp: 16 or older; Targ: under 16 plus 4 yr. age diff. <strong>Phipps (16)</strong></td>
</tr>
<tr>
<td>Sexual assault, 3rd degree</td>
<td></td>
</tr>
<tr>
<td>§ 61-8B-7</td>
<td>Gender neutral. Perp: 14 or older; Targ: 11 and under.</td>
</tr>
<tr>
<td>Sexual abuse, 1st degree</td>
<td></td>
</tr>
</tbody>
</table>

West Virginia
16
Adolescent Sex & the Workplace

<table>
<thead>
<tr>
<th>STATE CODE &amp; COMMENT</th>
<th>AGE OF CONSENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>W. VA. CODE § 61-8B-9</td>
<td>Gender neutral. Perp: 16 or older; Targ: under 16 plus 4 yr. age diff.</td>
</tr>
<tr>
<td>Sexual abuse, 3rd degree</td>
<td></td>
</tr>
<tr>
<td>Wisconsin 18</td>
<td></td>
</tr>
<tr>
<td>WIS. STAT. § 948.02</td>
<td>Gender neutral. Targ: under 13; Targ: under 16.</td>
</tr>
<tr>
<td>Sexual assault of a child</td>
<td></td>
</tr>
<tr>
<td>WIS. STAT. § 948.09</td>
<td>Gender neutral. Targ: 16 and over. **Phipps (16)</td>
</tr>
<tr>
<td>Sexual intercourse with a child</td>
<td></td>
</tr>
<tr>
<td>Wyoming 16</td>
<td></td>
</tr>
<tr>
<td>WYO. STAT. ANN. § 6-2-308</td>
<td>Gender neutral.</td>
</tr>
<tr>
<td>Criminality of conduct; Victim’s age</td>
<td></td>
</tr>
<tr>
<td>WYO. STAT. ANN. § 6-2-303</td>
<td>Gender neutral. Targ: under 12 plus 4 yr. age diff.</td>
</tr>
<tr>
<td>Sexual assault, 2nd degree</td>
<td></td>
</tr>
<tr>
<td>WYO. STAT. ANN. § 6-2-304</td>
<td>Gender neutral. Perp: 18 or older; Targ: under 14; Targ: under 16 plus 4 yr. age diff. **Phipps (16)</td>
</tr>
<tr>
<td>Sexual assault, 3rd degree</td>
<td></td>
</tr>
</tbody>
</table>
WILD DREAMERS: MEDITATIONS ON THE ADMISSIBILITY OF DREAM TALK

Louise Harmon*

INTRODUCTION .................................................................576
I. SOME EARLY THEORIES ABOUT DREAMS, O.J. SIMPSON, AND A FREUDIAN INTERLUDE .........................586
   A. Aristotle, Hobbes, and Some Premodern Dreamers ....586
   B. O.J. Simpson: Wild Dreamer Extraordinaire ...............591
   C. A Freudian Interlude: Simpson’s Fatal Obsession ............599
   D. Other Wilder, More Worrisome Theories About the Meaning of O.J. Simpson’s Dreams ..................605
II. THREE WILD DREAMERS FROM FAMILY COURT, A JUNGIAN INTERLUDE, AND THE DREAMING CULTURE OF THE IROQUOIS ...............................................................609
   A. Wild Dreamer #1 in Family Court Who Dreams of His Wife Exploding ..............................................609
   B. Wild Dreamer #2 in Family Court Who Only Dreams of Tomorrow .......................................................612
   C. Wild Dreamer #3 in Family Court Who Dreams of Her Own Murder (and Then Is Killed) .......................614
   D. A Jungian Interlude—Ayla’s Dream and the Archetype of the Murdering Father .................................620
   E. Concluding Ayla’s Story—The Dreaming Culture of the Iroquois ......................................................628
III. THE PENULTIMATE WILD DREAMER, CARTESIAN CONFUSION, ZHUANG ZHOU’S BUTTERFLY, AND A TIBETAN BUDDHIST INTERLUDE ........................................634
   A. Wild Dreamer Dreams He Has Murdered a Young Woman (and He Had Not) ........................................634
   B. Interlude: Tibetan Buddhism ........................................639

* Professor of Law, Jacob D. Fuchsberg Law Center, Touro College. I would like to thank Kristine M. Pizzo for her thorough and able research assistance, and Michael Estok, Zak Tomlinson, and Natasha S. Black for their excellent editing and for sharing their wild dreams with me.
INTRODUCTION

Dream talk. We all dream. We all talk. But we do not all talk about dreams. Every culture has its norms about the narration of dreams—where and when and how they may be talked about, and by whom. In modern western culture, there is no public forum for dream talk. Perhaps because we dream in sleep—alone, horizontal, and in the dark—we deem dreams to be private. Dreams are revealed in hushed tones at breakfast tables, across pillows, in therapists’ offices, in cars that drive through rainy nights. Dreams fall into no one’s public domain; they are on no one’s agenda for discussion.

Dream talk, and talk about dream talk, are rarely heard of in the law. The law belongs to the waking world, not to the world of sleep; it

1. The word “dream” is difficult to define. The first definition in Webster’s New Universal Unabridged Dictionary is “a sequence of sensations, images, thoughts, etc. passing through a sleeping person’s mind.” WEBSTER’S NEW UNIVERSAL UNABRIDGED DICTIONARY 555 (2d ed. 1986). In pointing out the problems in defining dreaming, one scholar rejects the view that there is one, and only one, kind of dream. He isolates a variety of kinds:

There do seem to be relatively distinct types of dreaming, each with its own line of development and potentially one-sided exaggeration—as permitted or perhaps demanded by the subductive clouding or single-mindedness of dreaming. There are relatively mundane dreams that seem to be based on mnemic consolidations and reorganizations; Freud-type relatively fantastic, pressure-discharge dreams, often based on complex rebuslike wordplay; dreams based on somatic states and illness; dreams based on aesthetically rich metaphor; dreams based on problem-solving and deep intuition (perhaps extrasensory?); lucid-control dreams; the varieties of nightmare; and a Jung-type archetypal-mythological form of dreaming.


2. For a fascinating discussion of “sleep talk,” or “vocalizations uttered during sleep” as opposed to the subject of this article which is “dream talk,” or narrations of dream experiences, see generally Deborah Rosenthal, Voices from Darkness: The Evidentiary Admissibility of Sleep Talk, 30 U.S.F. L. REV. 509 (1996).

3. Sleep does make a brief appearance in most criminal law texts under the rubric of the defense of somnambulism, a form of “automatism” in which “one who engages in what would otherwise be criminal conduct is not guilty of a crime if he does so in a state of unconsciousness or semi-consciousness.” WAYNE R. LAFAVE & AUSTIN W. SCOTT, JR., SUBSTANTIVE CRIMINAL LAW § 4.9, at 541 (1986). The leading case is Fain v. Commonwealth, 78 Ky. 183 (1879). In Fain, the defendant, while asleep in a hotel lobby, shot and killed a porter who was trying to wake him up. The Kentucky Court of Appeals reversed Fain’s murder conviction after the trial court refused to receive evidence that the defendant was a sleepwalker or to instruct the jury on an unconsciousness defense. Id. at 187 (quoting a nineteenth century treatise, Brown on the Medical Jurisprudence of Insanity: “Indeed, there are very many cases in which the confused thoughts of awakening consciousness have led to disastrous consequences. And this is to be accounted for by the fact that
Wild Dreamers

presupposes consciousness, social interaction, verticality. The sentences spoken and written by those who live with and in the law are crafted for the light. Besides, most of the law’s business is volitional. Dreams seem to happen to us.⁴ Often dreams make no sense; they defy rationality.

The law prides itself on rationality. The law also ascribes to the dominant epistemology of empiricism. As generations of philosophy undergraduate students have learned, empiricism demands that all knowledge derive from sensory perception. We can only claim to “know” that the desk the professor sits upon is “real” because we can see it, hear and feel our hands thumping on it, and, if we were so inclined, we could sniff it, lick it, and taste its woodiness or plasticity.⁵ The model is scientific; only after we have methodically gathered our sense data are we prepared to stick out our necks and make the ambitious claim that we “know” the table is really there. Dreams have no place in such an episteme. So-called “knowledge” that comes from dreams, from extrasensory perception, from visions, or from a spiritual experience is not really knowledge. One can make sentences about information received from a dream, a vision, or an intuition, but they constitute a different kind of claim—a claim of faith perhaps.

there is a state between sleeping and waking when the thoughts of the dreamer have as much reality as the facts he is assured of by his senses.”).

⁴. Our need for sleep has been scientifically verified, although scientists are uncertain about its biochemical purposes. See WILLIAM C. DEMENT, SOME MUST WATCH WHILE SOME MUST SLEEP 3 (1976). There are two kinds of sleep: REM (Rapid Eye Movement) sleep and non-REM sleep. BSTAN-'DZIN-RGYA-MTSHO, DALAI LAMA XIV, SLEEPING, DREAMING, AND DYING 29 (Francisco J. Varela ed. & narr., B. Alan Wallace & Thupten Jinpa trans., 1997) [hereinafter DALAI LAMA, SLEEPING, DREAMING, AND DYING]. Dreaming takes place during REM sleep. Id. In non-REM sleep, the brain becomes more silent, according to measurements of the global amount of brain electrical activity. Id. at 32. The heartbeat and respiration also slow down. Id. at 31. In REM sleep, there is rapid eye movement; muscle tone changes and significant cerebral activity are manifest. Id. at 31–32. The brain patterns in REM sleep correspond to the dream state, and when people are awakened from REM sleep, more than eighty percent report they were dreaming. Id. at 32. About twenty to twenty-five percent of the complete circadian cycle is spent in REM sleep. Id. at 33. REM sleep is universal in mammals. Id.

⁵. William James defines this theory of knowledge and use of logical reasoning as “rationalism.” His definition is as good as any:

Rationalism insists that all our beliefs ought ultimately to find for themselves articulate grounds . . . [:] (1) definitely statable abstract principles; (2) definite facts of sensation; (3) definite hypotheses based on such facts; and (4) definite inferences logically drawn. Vague impressions of something indefinable have no place in the rationalistic system, which on its positive side is surely a splendid intellectual tendency, for not only are all our philosophies fruits of it, but physical science (amongst other good things) is its result.

WILLIAM JAMES, THE VARIETIES OF RELIGIOUS EXPERIENCE 72 (1958). James goes on to point out that rationalism accounts for only a superficial part of man’s mental life, albeit the “part that has the prestige.” Id.
These principles of rational thought and of privileging only claims of knowledge based upon sensory perception are reflected in the rules of evidence. In a trial, the rules of relevance assume that jurors will use tools of logical inference to establish guilt or innocence. Another rule allows lay witnesses to testify only to matters about which they have personal knowledge—things they have perceived for themselves. We will allow percipients to come into court and claim to know such a thing, although the opposing party will have an opportunity on cross-examination to demonstrate that they did a lousy job looking, a lousy job listening. Conversely, we will not generally allow witnesses to testify to out-of-court observations of others—those statements constitute hearsay. The percipients would not be present in court, and not subject to cross-examination. There would be no way to test their credibility as lookers, listeners, competent narrators of all they have perceived.

6. Peter Murphy summarizes what William Twining has dubbed the “tenets of Optimistic Rationalism” that pervade the law of evidence and how we conduct our trials in the Anglo-American legal system:

The basic tenet of Optimistic Rationalism is that the drawing of rational inferences from relevant evidence is the only, or only known, method of correctly reconstructing past events. Twining analyzes the assumptions underlying this tenet, which include the epistemological assumptions: (1) that events and states of affairs occur independently of human observation; (2) that it is possible in principle to achieve present knowledge of past events; (3) that because evidence is typically incomplete, proof in judicial trials must be based on some degree of probability rather than absolute certainty; and (4) that a chain of inferential reasoning from evidence to hypothesis must depend on intermediate generalized assumptions about the causes and courses of human affairs.

Peter Murphy, Introduction to EVIDENCE, PROOF, AND FACTS 1, 19 n.16 (Peter Murphy ed., 2003).

7. The very concept of relevance is based upon a logical form of reasoning. Federal Rule of Evidence 401 defines relevance in this way: “Relevant evidence” means evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.” FED. R. EVID. 401. One evidence scholar explained some of the reasons we might call evidence relevant:

In some cases evidence seems relevant because we can construct a direct chain of reasoning from the evidence to some major, ultimate, or final hypothesis of concern. In other cases, however, we may have evidence from which we can form no such chains of reasoning, but the evidence seems relevant all the same because it bears upon the strength or weakness of links in the chains of reasoning we have already established from other evidence.


8. Federal Rule of Evidence 602 provides that a “witness may not testify to a matter unless evidence is introduced sufficient to support a finding that the witness has personal knowledge of the matter.” FED. R. EVID. 602.

9. See id. 801, 802.
Wild Dreamers

Rational thought and empiricism, with its firm commitment to sense data and rejection of the unseen world, have controlled the law.10

Does dream talk have any place in the law? And what about those of us who are wild dreamers, who try to live with and in the law?11 Where do we fit within the rigid confines of law’s rationality, with its insistence on the empirical episteme? How many lawyers would rather sleep than be awake?12 I would.13 But I do not sleep to escape what others call the

10. Historically, evidence law grew out of sixteenth and seventeenth century empirical philosophies. In her chapter on the development of the “beyond a reasonable doubt” standard, Barbara Shapiro writes about the intellectual traditions that the judges drew upon in their formation of the rules of evidence:

Thus the judges confronted twin sources of epistemological guidance. One was the English religious tradition, particularly the casuistical tradition, which sought a rational method of decision making in everyday life. The other was the scientific movement of Bacon, Boyle, and especially Locke and the empirical philosophers, who sought to establish scientific truth from the evidence they gathered.

BARBARA SHAPIRO, BEYOND REASONABLE DOUBT AND PROBABLE CAUSE: HISTORICAL PERSPECTIVES ON THE ANGLO-AMERICAN LAW OF EVIDENCE (1991), reprinted in EVIDENCE, PROOF, AND FACTS, supra note 6, at 327.

11. This article is only going to treat the dreams we have during sleep, although I suspect that many nocturnal wild dreamers are also wild daydreamers. I am confident there are many daydreamers in the legal profession. We daydream all the time, with one eye on our utterances to the outside world—in our law offices, courtrooms, and law classrooms—and the other eye turned inward. We remain quiet about those daydreams, worried that we will be deemed untrustworthy, not in touch with reality—a rudderless, oarless vessel drifting out to sea. Unlike sleep research and dreaming, there has been little research about daydreams, partly because there is no consensus about how to define a daydream. They can range from “fairly structured fantasies to the series of disconnected images that emerge during relaxed free association.” JAMES R. LEWIS, THE DREAM ENCYCLOPEDIA 59 (1995). While there seems to be some relationship between daydreams and regular dreams, “investigators in this area have hypothesized that somewhat different and perhaps independent cognitive processes are at work.” Id.

12. I am not alone in believing that sleep is not only a fundamental need, but a right. In the area of human rights, there has been pressure to expand the list of recognized rights due to every human being. Probably one of the longest lists is that proposed by Galtung and Wirak, of needs “that might be considered as important candidates on the world waiting list for processing into rights.” Galtung & Wirak, On the Relationship Between Human Rights and Human Needs, UNESCO Doc. SS-78/Conf.630/4, at 48 (1978), quoted in Philip Alston, Conjuring Up New Human Rights: A Proposal for Quality Control, 78 AM. J. INT’L L. 607, 610 (1984). At the top of the list: “the right to sleep.” Id.

For a discussion of sleep as symbolic speech, see Clark v. Community for Creative Non-Violence, 468 U.S. 288 (1984). In Clark, demonstrators set up a small tent city in Lafayette Park in Washington D.C. called “Reaganville” to demonstrate the plight of the homeless. Id. at 291–92. At issue were National Park Service regulations that forbade camping except in designated areas. Id. at 290–91. While the Court upheld the regulations, it also assumed, without deciding, that sleeping in “Reaganville” was symbolic speech. Id. at 293. The dissent was critical of what it perceived to be the majority’s evasiveness in not deciding that such activity is, or is not, speech. Id. at 301–02 (Marshall, J., dissenting). Chief Justice Burger stated in his concurrence that sleep was conduct, not speech. Id. at 300 (Burger, C.J., concurring).
real world. I sleep to enter the dream state, to establish contact with an unseen world that often reveals the truth to me. I have come to know many things from the dream state: about illnesses, my own and those of others; about past events; about hidden love. I have untied many knots in my dreams: solved problems, understood hearsay within hearsay, discovered a new direction in my thinking, and even found the words

13. I am also a Puritan, however, and try to keep my sleep to a medically acceptable minimum. There are norms for different species with respect to how much they sleep, and the bodily positions assumed for sleep. Humans tend to sleep about a quarter to a third of each twenty-four-hour period, and typically sleep lying down. DALAI LAMA, SLEEPING, DREAMING, AND DYING, supra note 4, at 34. Tigers sleep in trees. Id. Elephants sleep standing up, and only sleep on the average 3.2 hours of the day. Id. Rats sleep eighteen to twenty hours of the day. Id. Dolphins continue to swim since only half of their brain actually sleeps. Id. Biologists speculate that birds who migrate for many days sleep while they are flying. Id.

14. My dreams were most vivid and frequent during pregnancy. I kept having dreams of giving birth to a litter of kittens. Research has shown that pregnant women’s dreams are remarkably different from those of nonpregnant women. Pregnant women not only dream more, but the content of their dreams “is also unusual in that it is frequently vivid and rich in detail, bizarre, and often nightmarish.” Patricia Maybruck, Pregnancy and Dreams, in DREAME TIME AND DREAMWORK: DECODING THE LANGUAGE OF THE NIGHT 143, 143 (Stanley Krippner ed., 1990) [hereinafter DREAMTIME AND DREAMWORK]. Many pregnant women dream of giving birth to small mammals, and “small animals in their dreams [are] said to represent the unborn child.” Id. “Kittens, puppies, bunnies, and other small creatures [are] also typical fetal symbols.” Id.

There may also be a dream connection between the mother and her newborn. There have been many anecdotes about telepathic bonding between a mammalian mother and her infant. In mammals, when most of the postnatal time is spent in sleep, and a significant portion of the total sleep time in REM sleep, one researcher suggested that dream telepathy served “the important function of bridging the gap between a vulnerable dreaming organism exposed to threats and a protective adult, in most cases the mother, as fathers tend to leave the mothers long before the birth.” Jon Tolas & Montague Ullman, Extrasensory Communication and Dreams, in HANDBOOK OF DREAMS 168, 195 (Benjamin B. Wolman ed., 1979). Tolas and Ullman theorized that the dreaming young would locate the external threat of a predator, incorporate it into dream imagery, and telepathically communicate the danger to the mother who might be away, out of earshot, gathering food. Id.

15. Elias Howe was said to have invented the sewing machine in a dream. He was having trouble making a machine that would create a locking stitch. He dreamed that “he had been captured by natives who all carried spears which had holes through the blades.” VALERIE FRANCIS, ILLUSTRATED GUIDE TO DREAMS 56 (1995). “He realized that the position for the eye of the needle was at the point, and went on to produce the machine needles we know today.” Id.

16. Vivid dreaming might even make us more productive in our work. The psychologist Ellen Winner writes:

Strands of similarity are thus drawn between art, neurosis, play, daydreaming, nocturnal dreams, and ordinary forms of productive work. All of these forms of behavior are driven by a common ingredient: powerful wishes that cannot be fulfilled. Add to this ingredient a proclivity toward repression, and one becomes neurotic. But add to this same ingredient a proclivity toward sublimation, and one becomes an ambitious worker or, if one also possesses a mysterious ingredient called ‘genius,’ one becomes a creative artist or scientist.

Wild Dreamers

to say what during the day was unsayable. I travel, write poems, paint murals, design clothes. I sing arias, dance over lunar landscapes, fly over tree tops, encounter the dead and other inhabitants of the unseen world. Sometimes these nocturnal journeys sadden me; other times they fill me with joy. I do not know what these dreams do to or for me, but I know I need to have them. So I sleep.

17. Unfortunately, I always dream in words and not in numbers. In 1997, a former New Hampshire telephone operator and her husband shared a $66 million lotto prize when she played six numbers that appeared to her in a dream. After she woke up with the six numbers “dancing in her head,” she played the numbers and won the multistate Powerball lottery jackpot. Nation in Brief, THE ATLANTA J. & CONST., Dec. 24, 1997, at A9. “‘They came to me in a dream,’ Mary Sanderson said, ‘and, no, it didn’t have anything to do with the Psychic Network.’” Id.

18. I often travel to the house in which I grew up, both in dreams and in daydreams. Gaston Bachelard writes about the dreams and daydreams that we all have of the houses “we were born in.” GASTON BACHELARD, THE POETICS OF SPACE 16 (1964). When we visit the home of our past in dreams and daydreams, we reach the

real being of our childhood. . . . It is on the plane of the daydream and not on that of facts that childhood remains alive and poetically useful within us. Through this permanent childhood, we maintain the poetry of the past. To inhabit onerically the house we were born in means more than to inhabit it in memory; it means living in this house that is gone, the way we used to dream in it. . . . Beyond all the positive values of protection, the house we were born in becomes imbued with dream values which remain after the house is gone. . . . And we should not forget that these dream values communicate poetically from soul to soul. To read poetry is essentially to daydream.

Id. at 16–17.

19. Bad dreams are frequently experienced by victims of crimes, war, or other violence. See, e.g., Clemens v. Massanari, No. CV 00-6204-KI, 2001 WL 34043764, at *3 (D. Or. May 17, 2001) (where applicant who was beaten by a baseball bat reported nightmares and flashbacks as part of the symptoms of his Post Traumatic Stress Disorder (PTSD) claim); Wright v. Massanari, No. A-99-CA-808 AA, 2001 WL 694499, at *2 (W.D. Tex. May 10, 2001) (where applicant who was in a plane crash dreamed he saw blood and people with missing limbs as part of the symptoms of his PTSD claim); Curry v. West, No. 98-1980, 2000 WL 792316, at *2 (Vet. App. June 14, 2000) (where applicant had trouble sleeping due to nightmares and flashbacks as part of the symptoms of his provisional PTSD claim); State v. Catalano, No. M200103039CCAR3CD, 2003 WL 21877933, at *3 (Tenn. Crim. App. Nov. 24, 2003) (where child victims of aggravated sexual battery experienced nightmares); Fairfax County Fire & Rescue Dep't v. Mottram, 559 S.E.2d 698, 700 (Va. 2002) (where worker dreamed recurrent distressing dreams and suffered from sleep disturbance due to repeated exposure to traumatic stressors from his work as a paramedic supervisor as part of his workers’ compensation claim).

20. It is my firm belief that the world would be a better place if we all dreamed—and slept—more. This is true even for people who do not solve problems in their dreams. People now sleep twenty percent less than people who lived a hundred years ago, a decline that began with the widespread use of electric light. Nick Rufford, Bosses Wake Up to Benefit of Taking a Nap, TIMES NEWSPAPERS, July 2, 1995, at 2.7. Studies at Cornell University indicate that most business executives, for example, carry a “sleep debt” of between an hour and an hour-and-a-half from insufficient sleep at night that “impairs almost every mental faculty, including decision-making, creative problem-solving, attention span, concentration, memory and vocabulary.” Id. James Maas, professor of psychology at Cornell, offers a seminar, Asleep in the Fast Lane, to introduce daytime sleep therapy for business people in America and Asian-Pacific countries. Id.
Sometimes I enter the dreams of others, becoming an inhabitant in someone else’s unseen world. My visitations are occasionally reported. A friend will tell me that I was in her dream last night, and I am not surprised. Almost always the dreamer’s story is different from mine. In our sleep, we each craft a unique concatenation of events, our own unsturdy web of unlikely causal relations and bizarre backdrops, but our dreams have something in common: a knot with two connecting threads. I was there; she was there. I do not volunteer that I already know of our encounter. Revelations like that bother most people. They challenge widely held beliefs about what is real and what is not.

How could such a wild dreamer have become a lawyer? Even more confounding: how could such a wild dreamer have become a teacher of

21. Walt Whitman recorded visitations to the dreams of others in his poem The Sleepers. WALT WHITMAN, The Sleepers, in LEAVES OF GRASS (1855), reprinted in THE PORTABLE WALT WHITMAN 30, 115 (Mark Van Doren ed., 1945). Whitman claimed more than visitation, however; he also claimed union with all those who sleep and dream:

The earth recedes from me into the night,
I saw that it was beautiful . . . . and I see that what is not
the earth is beautiful.

I go from bedside to bedside . . . . I sleep close with
the other sleepers, each in turn;
I dream in my dream all the dreams of the other dreamers,
And I become the other dreamers.

Id. at 116.

22. Someone might label me as mentally ill. Some of my dreams are diagnostic. Sometimes in my dreams, and even when I am awake, I will get images of hot, violet-red pain in other people’s joints, particularly those from the hip down. I am also overly sensitive to the clogged kidneys of others that manifest themselves, again in both my waking and dreaming life, as a lime, green phosphorescent light emanating from around the neck and shoulders. Confessing to these kinds of images can be dangerous. In a discussion of healers and shamans, Holger Kalweit writes about how many wild dreamers are institutionalized:

For the most part, Western mediums and healers live in a culture that regards their experiences as unnatural, if not pathological, and mediumistic activities are looked at from the point of view of standard psychological norms and are therefore regarded as hallucinatory. . . . [I]n our culture the symbols of transformation are negative: they include hospitalization, schizophrenia, brain-wave tests, stupifying psychotropic drugs, and ostracism from society. How many unrecognized shamans, mediums, and saints fill the madhouses of rationalism? How many powers have been mangled and cut off during the long history of psychiatry? How many people has psychology reduced to mindless robots through its abasement of the psyche?


23. It is not so much that lawyers would reject my dream life per se, but they would not accept that anything learned during a dream could qualify as knowledge. William James would probably call dreams “mystical states of consciousness.” JAMES, supra note 5, at 292. Besides dreams’ ineffability, James points out, their other hallmark is their “noetic quality.” Id. at 293. James writes:
law? Teachers are authority figures; their job is to transmit knowledge, to inculcate the dominant values of the profession. How could a teacher operate out of a different—or at least augmented—epistemology from her students and colleagues? Were those who hired me just ignorant of my heresy? Or did they choose to hire a heretic—perhaps a court jester? Most of them are no longer alive, so perhaps it is not worth asking. Besides, being dead, they might be unable to reconstruct the narrow framework of thought that generated the conclusion, lo, these many years ago, that this wild dreamer should teach the law. Surely that must be one thing dying will do for you—it will broaden your horizons. 24 I am counting on that.

This article presents a mosaic of interrelated meditations on how judges have dealt with the admissibility of dream talk, interspersed with short digressions on the meaning of dreams from a variety of historical and cultural perspectives. In Part I, I begin with theories about dreams from Aristotle, Hobbes, and others. I then tell the story of O.J. Simpson, wild dreamer extraordinaire, interrupted by a Freudian interlude and a speculation about the theories that the jurors may have silently applied to Simpson’s dreams of killing his wife. In Part II, I present three wild dreamers from family court: a husband who dreams of his wife’s car exploding, a father who dreams about the future, and a six-year-old girl...

[M]ystical states seem to those who experience them to be also states of knowledge. They are states of insight into depths of truth unplumbed by the discursive intellect. They are illuminations, revelations, full of significance and importance, all inarticulate though they remain; and as a rule they carry with them a curious sense of authority for after-time. 24. James may have found that sense of authority “curious,” id., but “rational” thinkers—and most lawyers wish to be included in that category—find that sense of authority bogus.

24. There have been some fascinating cases about the establishment of testamentary charitable trusts where the dead have tried to ensure that their worldly possessions go to some worthy cause that benefits them. The law of trusts and estates is also firmly grounded in the empiricist episteme. An English case demonstrates the law’s hostility to gifts to any organization that ascribes to belief in paranormal phenomena. In Beatty v. London Spiritualistic Alliance, 1 Ch. 237 (1923), the decedent left a legacy to set up a trust to establish a college for the training of mediums. Id. at 239–40. Under the law, in order to establish such a trust, it had to be “a good charitable gift,” namely, one that would benefit the public. Id. at 240. Justice Russell held the gift to be invalid:

If a testator by stating or indicating his view that a trust is beneficial to the public can establish that fact beyond question, trusts might be established in perpetuity for the promotion of all kinds of fantastic (though not unlawful) objects, of which the training of poodles to dance might be a mild example.

Id. at 242. The court expressly reserved the right to determine whether a college for mediums was foolish. I, for one, would love to leave some money to establish a college for dancing poodles. For a less demeaning assessment of spiritualism and the training of mediums, see In re Lockwood’s Estate, 25 A.2d 168 (Pa. 1942). See also Mary Kay Lundwall, Inconsistency and Uncertainty in the Charitable Purposes Doctrine, 41 WAYNE L. REV. 1341 (1995).
who dreams of her own murder by her father. I explore how this girl would have fared under Carl Jung’s theory of dreaming, or in the dreaming culture of the seventeenth century Iroquois. In Part III, I tell the story of the penultimate wild dreamer who dreamed he had killed a young woman (but had not) and ponder his confusion from the point of view of Descartes and the ancient Chinese philosopher Zhuang Zhou, ending with an interlude about what dreams mean in Tibetan Buddhism. In Part IV, I present the tale of the ultimate wild dreamer—a man who dreamed someone else had killed a young woman and ended up convicted of that crime.

Here is my central observation: how a judge will rule on the admissibility of dream talk will depend upon his or her theory about the meaning of dreams. That theory in turn will depend upon his or her metaphysics and episteme. Here is my conclusion: given the dominance of rationalism and the empirical episteme, evidence of dream talk should not be admissible.

My conclusion may seem inconsistent with my avowed experience as a wild dreamer. How could someone who believes that knowledge can be obtained from dreams argue against their admissibility? As a lawyer, I too have been indoctrinated in rationalism and the dominant, empirical episteme. I might be willing to accept a dream as a source of knowledge for myself, in my own private waking world, but I am not willing to send someone to prison, or to decide issues of child custody, on the basis of dream talk. Dream talk is too speculative. Because dreams are privately experienced in our culture, we have developed no criteria by which to judge their credibility. There are no shared assumptions about the meaning of dreams, no accepted principles about how to interpret them, no social norms about whether any given dream reveals the truth or is just foolishness. Under the empirical episteme, there is some agreement about what constitutes a valid claim of knowledge, and what does not. We are equipped to test the reliability of sentences that start out: I saw this, I heard that, I felt this, or I smelled that. We are not equipped to test the reliability of sentences that start out: I dreamed this or I dreamed that. We share no assumptions about the dream world; we have no common language to make sense of each other’s dream talk.

Even more worrisome: no one can predict what theories about the meaning of dreams the jurors might follow. They sit silently in judgment. There is no way to detect what their metaphysical and epistemological assumptions might be. Indeed, it is precisely my status as a wild dreamer that compels me to argue against the admissibility of
Wild Dreamers
dream talk. I am all too familiar with how easy it is to mask one’s competing world view beneath a veneer of rationality—even from one’s self. Too much authority might be granted to dream talk. Someone might think that God was talking, for example, or that the direction of time was temporarily reversed. For all anyone knows, the jurors’ assumptions about the source and meaning of dreams may be as wild as mine—or wilder. I know a man, a staunch member of the community, a sweet, bald, middle-management kind of guy, who believes that aliens are not only walking among us, but actively intruding in our dreams. He worries that through his dreams, his waking life is being micro-managed by a busy body from a distant corner of the universe. In my private life, I am willing to concede that any of those things might be happening, but not in the courtroom. Not with someone else’s life at stake.

How did I become interested in the admissibility of dream talk into evidence? During the O.J. Simpson trial, Judge Ito admitted a statement by Simpson that he had dreams of killing his wife, and I was bothered by that. I am still bothered by that. A lot. One day, while daydreaming about the Simpson trial, I discovered that on occasion the dream world interjects itself into the rational, empirical, waking world of the law. More often than not, these intrusions occurred in criminal prosecutions for murder. There were other kinds of cases as well, mostly from family court, in which dreams were offered to prove the unfitness of a parent, or to show the dreamer’s damaged state of mind. I became intrigued with the clash of epistemes.

As I researched the subject of the admissibility of dreams, I also began to wonder if there was any role for the wild dreamer to play in the legal profession and its academy. A hidden agenda developed, although as I reveal it, I suppose it loses its essence of subterfuge. Does the wild dreamer who lives with and inside the law have anything to contribute? And the wild dreamer law professor? Those questions are close to the surface of this article, still under water, but getting enough sunlight to know that they are there.

Enough introduction. Let me begin with Aristotle and Hobbes, and some premodern theories of dreaming, just to get us started. Then I will tell you the story of O.J. Simpson who was a wild dreamer extraordinaire. Freud will erupt in the middle of that story. Do not despair at the article’s lack of linearity. Just pretend you are in a dream.

25. See infra Part I.A.
I. SOME EARLY THEORIES ABOUT DREAMS, O.J. SIMPSON, AND A FREUDIAN INTERLUDE

A. Aristotle, Hobbes, and Some Premodern Dreamers

I begin with the mundane, a characterization that Aristotle, who always had his feet planted firmly on the ground, would not have quarreled with. Aristotle’s theory about dreams and their meaning—or lack thereof—is the least ambitious, and consequently, the easiest one to grasp. I also suspect that for those readers who are not psychologically inclined, Aristotle’s theory about dreams will still resonate. It sounds remarkably modern, and for that reason, belongs more to the empirical episteme than to what I shall call the premodern or medieval.

Dreams, Aristotle believed, were the result of impressions that had been experienced in the physical world, and were re-experienced in the

---

26. Aristotle departed from the earlier Greek view of dreams in which it was believed that a god would actually make a visit to a sleeping person in some recognizable form. These dreams were often the result of incubation—“the practice of seeking dreams for specific purposes.” LEWIS, supra note 11, at 4. People went to special temples devoted to the god Aesculapius, the most popular healing divinity in the Hellenistic world. Id. There they would fast, engage in rituals, and sleep, with the intention of receiving a healing dream. Id. Other gods made visitations in dreams as well. In The Iliad, Zeus sent dreams to chosen people (the recipient was always a male). See HOMER, THE IliAD book 2, at 100, lines 33–34 (Robert Fagles trans., 1991). In The Odyssey, Athena sent dreams to female recipients (for example, Athena sent a phantom to Penelope’s bedroom to tell her at the "gate of dreams" about how Telemachus, her son, was alive and well). See HOMER, THE ODYSSEY book 4, at 150–51, lines 910–45 (Robert Fagles trans., 1996). As one scholar put it:

Dreams were considered to be objective facts, things that happened to you. The Greek was ‘visited’ by a dream (episkopein), at best he ‘saw’ a dream (enypnion idein). They would never have dreamed of saying as the French do nowadays, ‘J’ai fait un rêve,’ or the Italians, ‘Ho fatto un sogno.’

Carl Alfred Meier, The Dream in Ancient Greece and Its Use in Temple Cures (Incubation), in THE DREAM AND HUMAN SOCIETIES 303, 308 (G.E. Von Grunebaum & Roger Caillois eds., 1966). This purely religious attitude towards dreams was replaced by the more rationalistic tradition exemplified by Aristotle.

27. I confess to using the word “modern” very loosely. I do not mean to imply that the medieval period was the only premodern historical age; obviously, for example, Aristotle belonged to classical Greece, having lived sometime in the fourth century B.C.E., although his ideas often seem more “modern.” “The word modern, first recorded in 1585 in the sense ‘of present or recent times,’ has traveled through the centuries designating things that inevitably must become old-fashioned as the word itself goes on to the next modern thing.” THE AMERICAN HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE (4th ed. 2000), http://www.bartleby.com/61/28/M0362800.html. I am using it more in that broader sense, although historically, the empirical episteme became dominant within the historical period typically dubbed “modern.” At any rate, I acknowledge sloppiness of usage here in using the terms “premodern” and “medieval” synonymously. Some of my famous premodern wild dreamers were medieval, but others belonged to much earlier periods, and some of us are still around.
Wild Dreamers

dream in a random fashion. He pointed out that people usually dream at night of the same experiences that occurred earlier in the day; similarly, a dream at night might cause someone to get up in the morning and repeat the same experience. Since the soul is more sensitive at night, Aristotle argued, it is more likely to register sensations from the outside, magnifying their degree and changing their quality:

This is plain in what often happens during sleep; for example, dreamers fancy that they are affected by thunder and lightning, when in fact there are only faint ringings in their ears; or that they are enjoying honey or other sweet savours, when only a tiny drop of phlegm is flowing down [the oesophagus]; or that they are walking through fire, and feeling intense heat, when there is only a slight warmth affecting certain parts of the body. When they are awakened, these things appear to them in this their true character.

For Aristotle, the dream was nothing more than the result of somatic stimuli. There was no divine agent; God had nothing to do with it.

As they ushered in the modern age, the empiricists of the sixteenth and seventeenth centuries expanded upon this Aristotelian belief that dreams were merely a biological response to external stimuli. Hobbes was exemplary, expressing a dogged dualism that was the hallmark of western philosophy. He had no trouble distinguishing the honey of his

29. Id.
30. Id.
31. Id. Aristotle embraced many of the principles that later belonged to empirical science and rejected Platonic metaphysics. In De Anima, for example, he developed a theory of the soul that treated the soul and the body as constituting a single entity, standing to each other in the relation of form to matter. ARISTOTLE, DE ANIMA 156–57 (Hugh Lawson-Tancred trans., 1986). Aristotle believed that the waking state was determined by the activation of the common-sense faculty, and the sleeping state by its inactivity. The illusion of “sense-perception” in sleep is due to “the improper functioning of the senses, which frees the way for the forming of dreams, without correction by judgment or evaluation.” LEWIS, supra note 11, at 19. Aristotle wrote about dreams in De Somno et Vigilia, De Insomniis, and De Divinatione per Somnum. Id. at 18.
33. Hobbes went even further than Aristotle, assuming that all dreams were the result of physical factors:

And seeing dreames are caused by the distemper of some of the inward parts of the Body; divers distempers must needs cause different Dreams. And hence it is, that lying cold breedeth
dreams from the phlegm that flowed down his esophagus. When Hobbes was awake, he was not dreaming, and what he experienced was real. When Hobbes was asleep, he was dreaming, and what he experienced was not real. The dreaming state was just the reverse of the waking state. “And that as Anger causeth heat in some parts of the Body, when we are awake,” Hobbes argued,

so when we sleep, the over heating of the same parts causeth Anger, and raiseth up in the brain the Imagination of an Enemy . . . . In summe, our Dreams are the reverse of our waking Imaginations; The motion when we are awake, beginning at one end; and when we Dream, at another.34

Aristotle and Hobbes both denied that dreams could have a divine origin, or that they could be interpreted through the use of supernatural skills. Their beliefs were in stark contrast to the premodern or medieval tradition with its belief in an ultimate, transcendent reality, the worldly reality being only its sign or symbol.35 In the premodern era, the universe and the individual life were centered on God; man’s purpose was outside of himself. This meant that the source and rationale of all the incidents of man’s life, including dreaming, were divine.36 Thus, under Jewish and Christian belief systems, the dream was often deemed to be a vehicle for divine/human encounters.37 God appeared to a number of patriarchs, including Abraham, Jacob, and Joseph; the dreams were usually prophetic and often exhorted the dreamer and his people to

Dreams of Fear, and raiseth the thought and Image of some fearfull object (the motion from the brain to the inner parts, and from the inner parts to the Brain being reciprocall[]).


34. Id.
36. In premodern belief systems, “the dream is seen as possessed of cognitive force in regard to otherwise inaccessible sectors of objective reality, especially such as the future and the Hereafter, or, more generally, truths bearing on man’s relation to the divine.” Meier, supra note 26, at 6.
37. The Book of Daniel consists largely of a series of dreams and visions; it is the most complete treatment of dreams in the Talmud. Daniel interpreted the dreams of King Nebuchadnezzar who was afflicted by insomnia brought on by his awareness of portentous dreams that he could not remember. After consulting a number of Babylonian soothsayers who were unable to help him (and who paid for their failure by execution), Daniel sought assistance from God who revealed the meaning of the king’s dreams. One dream prophesized that his kingdom would one day be divided; another exhorted the king not to assume he was all powerful, a warning Nebuchadnezzar did not heed, paying for his hubris by being afflicted with a strange psychosis that lasted seven years. See generally Book of Daniel.
Wild Dreamers

take a particular course of action. The New Testament has a number of dreams that were divinely inspired—for example, the Lord’s visitation to Joseph in a dream, telling him to take Mary for his wife, since “that which is conceived in her is of the Holy Ghost.” Both the Jewish and Christian traditions regarded the dream as highly symbolic, requiring the assistance of a skilled interpreter to decode its meaning; almost always the dream was a revelation from God directing the course of human affairs.

38. In biblical accounts, the dream has sometimes served as a means of encountering God directly. The most prolific dreamers in the Talmud were the patriarchs, in particular Abraham, Jacob, and his eleventh son, Joseph. All of their dreams were forms of divine communication; most were direct revelations from God about his involvement in human affairs. Abraham, the ancestor of the Jews, was visited by God in a deep sleep. He was told the prophecy of the enslavement of the Jews for four hundred years and made a covenant to eventually save them. Genesis 15:12–16. The patriarch Jacob had the famous dream of “Jacob’s ladder,” and his son, Joseph, was not only a famous prophetic dreamer, but a famous dream interpreter. See id. 28:11–16. When Joseph interpreted the Pharaoh’s dream to foretell seven years of hunger and famine in Egypt, he recommended a plan to save the nation from the famine, and the Pharaoh was so impressed, he made Joseph prime minister. When Joseph’s family came to Egypt to buy grain, they honored him, according to earlier dreams that he had experienced, in which not only his family, but “the sun and the moon and the eleven stars made abeisance to me.” Id. 37:9.

39. Matthew 1:20. This dream announced to Joseph the conception of Christ. There were also three dreams connected to the flight into Egypt: the dream of the wise men who were warned not to return to Herod; the dream in which one of God’s agents, an angel, tells Joseph to go into Egypt to escape Herod; and after Herod’s death, the dream in which the agent again tells Joseph to return to the land of Israel. LEWIS, supra note 11, at 33.

40. The Islamic tradition also holds that dreams are revelations from God. LEWIS, supra note 11, at 131–32. Islam is essentially “a prophetic religion that is based on a series of divine revelations given to the prophet Muhammad through an angel during the latter part of his life,” sometime around 610 to 632 C.E. Id. at 131. Some of Muhammad’s spiritual instruction came through the medium of dreams. Id. In Islam, the Koran spells out a classification of dreams to Joseph. They are of three kinds: dreams of glad tidings from God; dreams of warning from the devil; and dreams that originate in the self. Von Grunebaum, supra note 35, at 7–8 (citing Abdalghani ad-Nabulusi, Ta’ir al-anidm fi ta’ir al-manam, 1, at 3–4 (1316)). Dreams could also, among other things, constitute a private prophecy, bear on politics, or be used to elucidate theological doctrine. See id. at 11–21.

In the Hindu tradition, the sacred literature is full of references to dreams, and most of them suggest that dreams are somehow connected to the supernatural. In the Rig Veda, for example, there are hymns for avoiding the evil effects of dreams that were said to cause disease and ill fortune. KRISHNA DAS GUPTA, THE SHADOW WORLD 52–53 (1971) (citing several Hymns of the Rig Veda). In the Brihad-Aranyaka Upanishad, there were two theories about dreams. Brihad-Aranyaka, Fourth Adhyaya, 3.7–34, in THE THIRTEEN PRINCIPAL UPANISHADS 134–39 (Robert Ernest Hume trans., Oxford Univ. Press 2d ed. 1995). One held that the soul took information from the world and constructed objects according to its own choice and desire. The other, which is very important as being connected with the idea of the immortality of the soul, states that during sleep the soul gets away from the body and experiences things which were distant from the sleeper. This theory implies that the soul roaming at a distance from the body may not find its way back into it and this has been regarded as an evil, very difficult to cure.
During the medieval period, mistranslations by St. Jerome of key biblical passages led to the notion that demons might visit someone during a dream. The central idea was that Satan’s minions were sent out to sleeping souls to lure them into temptation. Benedict Peterius, a sixteenth century Jesuit priest, wrote: “The devil is most always implicated in dreams, filling the minds of men with poisonous superstition and not only uselessly deluding but perniciously deceiving them.”

St. Hildegard of Bingen, a tenth century nun, was a wild dreamer and categorized her dreams as either prophetic or simply mundane, but she too “warned people to beware of demons masquerading as divine beings.” Perhaps nowhere was this suspicion of dreams more clearly demonstrated than in the notion of the incubi and succubi, particularly during the Inquisition. These were demons who took the form of handsome men and women who seduced the dreamer in his or her sleep. The incubi and succubi explained how otherwise celibate people—perhaps those who had given themselves up to God—could awaken in a state of sexual arousal with impunity—a medieval version of “the devil made me do it.”

Historically, it was too late for anyone in the courtroom in O.J. Simpson’s murder trial to seriously argue that his dream of killing his wife was the work of demons. Neither was the biologically based theory of dreaming on anyone’s mind. Rather, the Simpson case was tried about a hundred years after Sigmund Freud published The Interpretation of Dreams. As the prosecutors sought to introduce evidence of Simpson’s dream talk, they were counting on the fact that everyone in the

DAS GUPTA, supra, at 53.

41. LEWIS, supra note 11, at 158. Apparently St. Jerome was preparing a Latin translation of the Bible during the fifth century, and he consistently mistranslated a Hebrew word, resulting in a prohibition against any kind of dream work in the sacred text. As a result, dreams were “categorized with witchcraft for more than a thousand years.” Wilse B. Webb, Historical Perspectives: From Aristotle to Calvin Hall, in DREAMTIME AND DREAMWORK, supra note 14, at 175, 179.

42. LEWIS, supra note 11, at 158.

43. FRANCIS, supra note 15, at 15.

44. LEWIS, supra note 11, at 123.

45. Id. at 158. An incubus was male, a succubus female, and they worked in concert. Incubi were sterile, but could still impregnate women with seed taken by succubi from men—a belief that was sometimes used to explain pregnancies from secret liaisons. Id. at 158.

46. Id.

courtroom knew about Freud’s theory and how Simpson’s dreams revealed what was on his unconscious mind.

Not only that: the prosecutors may have been counting on the fact that some of the jurors might not be wholly rational in their thought. Their indoctrination in the dominant episteme might not have taken hold. As one scholar put it, certain contemporary attitudes towards dreams are incompatible with our concepts of infinite universe and of causation, in short with the major assumptions of both our cosmology and our epistemology. . . . Like an erratic bloc surviving from a past geological age the popular quest for material guidance from dreams, fed by curiosities and anxieties, continues on the medieval pattern.48

As Simpson’s dream talk was revealed, most of the jurors were probably steered by the prosecution to give it meaning under Freud’s theory of dreaming. But some of the jurors may have continued on the medieval pattern. Either way, how the jurors interpreted Simpson’s dreams posed grave danger to his defense.

But let me tell you first about O.J. Simpson, wild dreamer extraordinaire.

B. O.J. Simpson: Wild Dreamer Extraordinaire

O.J. Simpson was a dreaming man. The admissibility of evidence about Simpson’s dreams was at issue in his trial for the brutal murders of his wife, Nicole Brown Simpson, and her friend, Ron Goldman, by multiple stabbings, slashings, and cuttings.49 Ron Shipp, a former Los Angeles police officer and would-be actor, was a friend of O.J. Simpson.50 On the evening of June 13, 1994, the day after the murders, Simpson asked Shipp to accompany him to his bedroom to ask questions about the police investigation of the crime.51 According to Shipp’s statement to the police, Simpson first inquired about how long it took for DNA evidence to come back:

48. Von Grunebaum, supra note 40, at 3, 6.
49. Record, Motion re Admissibility of Statements of Ronald Shipp, Examination of Ronald Shipp, People v. Simpson (L.A. Super. Ct. 1995) (No. BA097211), 1995 WL 37667, at *1–*13 [hereinafter Motion re Admissibility]; Record, Motion re Admissibility of Statements of Ronald Shipp (Resumed), Examination of Ronald Shipp, Simpson (No. BA097211), 1995 WL 39921, at *1–*9 [hereinafter Resumed Motion re Admissibility]. I am grateful to Professor Martin Schwartz for alerting me to the dream talk issue in the Simpson case.
50. Motion re Admissibility, supra note 49, at *18.
[Shipp:] Yeah, it just came out of the blue . . . I mean he was just telling me regular O.J. talk and this just came out of the blue[.] He says, “How long does it take for D.N.A. evidence to come back?” I honestly have no idea, but I told him [“]two months.” And uhm, then uhm, I don’t know [we’re] watching T.V. for a while and then he said, oh, uhm, let me back up. Way before he brought up the D.N.A[.], he was telling me about this interview and he did not name the detective[,] he did not name you guys, but he said, “I was interviewed by detectives and they asked me to take a lie detector test.” And I replied, “Well, what did you say?” And he kind of chuckled and he says, “Hey, to be truthful Ron, Man, I have had a lot of dreams about killing her.” And he says, “I really don’t know about taking that thing.” He did not say [“]I won’t take it.”[”] He said, “I really don’t know about taking it.”

[Detective Vannatter:] Did he elaborate on these dreams?

[Shipp:] No. And I really did not ask him, because I said, “I am in shock man.” I just did not know what to say to this man . . . and at this point in time, it’s not personal. I mean from being a police officer, I just knew my personal belief, that he was guilty by asking these questions then.52

For a month, Ron Shipp did not reveal this conversation to anyone but his wife.53 When police detectives interviewed him, he did not tell them about Simpson’s statement that he had dreamed of killing his wife; neither did he subsequently tell the prosecutors.54 In Shipp’s later testimony at trial, he explained his failure to reveal it as a mixture of concern for Simpson (“I mean I loved this man for 26 years”),55 shock, and a desire to distance himself, wanting “nothing to do with any of this

52. People’s Points and Authorities in Support of Admitting Defendant’s Statement to Ron Shipp, Simpson (No. BA097211), 1995 WL 39922, at *3–*4 [hereinafter Points & Authorities].
53. Record, Examination of Ronald Shipp (Resumed), Simpson (No. BA097211), 1995 WL 37668, at *8. Similarly, when he was interviewed a couple of times by the defense team, Shipp was asked to tell them the worst things he knew about O.J. Simpson, and he still did not reveal the conversation. Id. at *12. While Shipp agreed that “this statement about the supposed dream is a pretty bad thing about Mr. Simpson,” and that he had lied about the statement about the dream, he felt both guilty towards the victims and loyal to the defendant. Id. “I really did not want to be really involved in all of this and I didn’t want to be going down as a person to nail O.J.” Id.
54. Id. at *8.
55. Id.
Wild Dreamers

at the time. I just—I was just still thinking it was a dream like most of America."  

A month after the murders, Ron Shipp was interviewed by Sheila Weller for her book *Raging Heart*, and consented to her relaying the conversation, as long as she referred to him by the pseudonym Leo. Shipp revealed the conversation to Weller, and later to the police and the jury: “for my conscience and my peace of mind. I will not have the blood of Nicole on Ron Shipp. I can sleep at night unlike a lot of others.” The conversation with Simpson about his dreams of killing Nicole was “eating” him up, and he initially talked to Weller, on the condition of anonymity, hoping to “exercise [sic] this pain” from his body.  

When Ron Shipp was set to testify, Simpson’s defense team hotly contested the admissibility of the entire statement, including both Simpson’s reluctance to take the polygraph test and his dreams of killing Nicole. Also contested was Simpson’s puzzling chuckle sandwiched between Simpson’s answer to Shipp’s question about how Simpson had responded when detectives asked him to take the lie detector test and Simpson’s “to-tell-the-truth” seven-word admission that “I’ve had a lot of dreams about killing Nicole.”

---

56. *Id.*
57. *Id.* at *15–*16.
58. *Id.* at *18.
59. *Id.* at *16. Shipp admitted that once Weller’s book revealed the conversation about the dreams between “Leo” and Simpson, he believed someone from the police or prosecution team would investigate the source of that conversation. *Id.* at *14. Fearing that he might have to perjure himself (“I did not want to sit up here and lie. Like I said, I’ve never lied on the stand ever in my life and I’m not starting now,” *id.* at *18.), Shipp contacted law-enforcement authorities himself a day or two before the book came out. *Id.* at *17.
60. See Motion re Admissibility, *supra* note 49, at *1–*13.
61. What was Simpson chuckling about, defense counsel argued? Simpson’s counsel said:
   Is he chuckling because the whole idea of a polygraph was funny because he is innocent and of course he passed? Is he chuckling because it is a silly idea because he is innocent and he throws in just by happenstance the sentence, ‘Well, you know, I’ve had dreams about killing her?’ Is he chuckling because the whole idea of having dreams is funny to him as well? Who knows.

*Id.* at *13.

There was some confusion about which of the statements the chuckle related to: Simpson’s reluctance to take the polygraph test, or Simpson’s statement about dreams of killing Nicole. *Id.* at *10–*11. The prosecution yielded to the admission of the chuckle, as long as it reflected upon the statement about dreaming of murdering his wife. *Id.* About the defense position that the chuckle related to Simpson’s “flippant attitude about polygraphs,” the prosecutor argued:

What on earth does that begin to say about the context of the statement that we are concerned with, about dreaming about the murder of his wife? That is the statement we are interested in
The prosecution yielded on the statement about the polygraph test but insisted on the admission of the dream talk, arguing that the dream was relevant to Simpson’s state of mind. Building upon its contention that Simpson was a wife-batterer who went too far, the prosecution argued that the statement was circumstantial evidence of a—what you could call a fatal obsession with this woman. And here it is almost as if he is saying I have a fatal obsession with Nicole except people don’t talk like that, and the way that he expresses it is he says I have been having a lot of dreams about killing her, and to me that seems like it is so obviously powerful compelling evidence of this fatal obsession with this woman that it is almost difficult for me to—to think of—to articulate, to put it into words. It is one of those things that almost seems to leap out.

The defense obviously quarreled with the prosecution over this claim of relevance. While the Dream Team was not prepared to produce expert testimony at the hearing on the statement’s admissibility, it argued that “whenever there are statements of dreams or questions of dreams, it is not necessarily indicative of the state of mind of the speaker.”

Putting into context... [H]e should have argued that somehow the chuckle reflects upon that statement, but to the extent that it does, the chuckle can come in.

Id. at *11.

62. The prosecution yielded because they had to. According to the U.S. Supreme Court, “[T]here is simply no consensus that polygraph evidence is reliable. To this day, the scientific community remains extremely polarized about the reliability of polygraph techniques.” United States v. Scheffer, 523 U.S. 303, 309 (1998). One fear is that by “its very nature, polygraph evidence may diminish the jury’s role in making credibility determinations.” Id. at 313. Moreover, such evidence is explicitly excluded under the California Evidence Code. See CAL. EVID. CODE § 351.1 (West 1995). In the prosecution’s words, “Since references to polygraphs are not admissible in evidence... the brief statements of the defendant that the detectives asked him to take a lie detector test and that he ‘didn’t know about taking one’ should be omitted.” Points & Authorities, supra note 52, at *3.

63. The prosecution got around the hearsay rule by arguing that Simpson’s statement was admissible under California Evidence Code Section 1220 as an admission of a party. Points & Authorities, supra note 52, at *2. Such statements need not be incriminating. See id. (citing People v. Aho, 199 Cal. Rptr. 671, 674 (1984); People v. Perkins, 180 Cal. Rptr. 763, 768 (1982)).

64. Id. at *3.

65. Motion re Admissibility, supra note 49, at *11.

66. The Dream Team was the name given by the press to Simpson’s defense team. It consisted of a cadre of famous lawyers, including Johnnie Cochran, F. Lee Bailey, Robert Shapiro, Alan Dershowitz, Barry Scheck, and Peter Neufeld. For a criticism of the term “Dream Team,” and a description of how the team was put together, see VINCENT BUGLIOSI, OUTRAGE: THE FIVE REASONS WHY O.J. SIMPSON GOT AWAY WITH MURDER 34–42 (1996).

was no “probative nexus between someone saying that I dreamt about something, assuming for the sake of my statement that the statement was made, and that it means something that is relevant to the trier of fact.”

The defense argued that O.J. Simpson was a distraught, weary, and unreliable narrator of the things of which he had dreamed, in the sleep he was not getting. To let the jury hear these seven words was highly prejudicial—too high a price to pay for evidence of negligible probative value. The defense threatened to bring out a cadre of experts:

[These experts will] talk about dreams and particularly in the context when this man is told that very day that a woman with whom he had been involved for seventeen years had just died. . . he is told on the telephone that this happens and he rushes back and he is consumed with all of this grief and all of this uncertainty and all of this emotion and he is tired because he didn’t sleep going to Chicago and didn’t sleep coming back and there are these family members that are also grieving with him. There is an entire context, your honor, that it would be unfair to require that we simply examine under a microscope six or seven different words without appreciating the entire context in which it is offered.

Judge Ito focused on the polygraph references, but expressed some curiosity about what inferences a jury might draw from Simpson’s dream statement that “I have had a lot of dreams about killing her.” He indicated that “it is a little more indirect as a thought process. So the

68. Id.

69. Alan Dershowitz, O.J. Simpson’s “hired constitutional guardian,” stated in a telephone interview that Judge Ito’s ruling admitting Simpson’s statement that he had dreams of killing Nicole constituted “a very significant issue on appeal here.” Brian McGory, Attorney Sees Grounds To Appeal Simpson Case, THE BOSTON GLOBE, Feb. 3, 1995, at 3. “We have looked at 10 cases around the country, all of which say dreams are not admissible. They say they’re prejudicial, because dreams do not reflect reality.” Id. Obviously, since Simpson was acquitted, there was no need for an appeal.

70. When proposed evidence is of a novel type, subject to a significant debate within the scientific community, most jurisdictions allow a hearing to determine whether to admit this kind of expert testimony. In California, it was called a “Kelly-Frye hearing” at the time of the Simpson trial. See CAL. EVID. CODE § 801 (West 1995). Oddly enough, the defense did not request such a hearing regarding the admission of dreams, and Judge Ito did not order one on his own. Although the defense implied that there would be such a hearing, see Motion re Admissibility, supra note 49, at *9 (“Therefore, there is going to be a great deal of scientific testimony that I will have to try to obtain to put on people that will be able to talk about dreams.”), it seems from the record that there never was such a hearing on the scientific significance of dreams—an oversight in my mind.

71. Motion re Admissibility, supra note 49, at *9.

72. Id. at *11.
question then becomes what is the probative value of, you know, putting your head on the pillow and having dreams about killing your wife?"73

Ito asked the parties if either side had “found cases that deal with admissibility of dreams.”74 The prosecution, the proponent of the evidence, confessed “quite frankly we didn’t even research that, your honor, because it seemed like our thought process was much the same as the court’s and that we were more concerned about the issue of the polygraph and redacting that out.”75 The prosecution argued that the jury should be able to use “their collective life experience and their common sense for the purposes of interpreting evidence . . . . [T]hat is all we are

73. Id.
74. Id. at *12. Despite this request, it became clear from the transcript that Ito planned to admit the statements about Simpson’s dreams when he began to inquire into the mechanics of how the prosecution was going to question the witness, redacting out the references to the polygraph test, and including the statement about the defendant’s dreams. Id. (“All right. Mr. Goldberg, how would you propose to present this, knowing the mine fields that are apparent on either side of this sentence?”).
75. Id. Judge Ito was the only participant at the hearing who cited to any case law about the admissibility of dream talk. Resumed Motion re Admissibility, supra note 49, at *3. Ito urged both parties to read State v. White, 156 S.E.2d 721, 722 (N.C. 1967), a murder case in which the defendant, Bobby Joe White, was prosecuted for the shooting of his good friend, Bud Brown, in a poolroom brawl. Resumed Motion re Admissibility, supra note 49, at *3. In White, a witness, Booker T. Brown, testified that he had worked with the defendant the day before Bud Brown had been shot, and when asked about a conversation with the defendant on that day, he said, “‘We were riding to work together every day. That Thursday morning we were going to work, he said he dreamed that he had shot Bud.’” White, 156 S.E.2d at 723. The defendant denied ever having made such a statement, and argued it was error to admit it since the jurors were not competent to interpret a dream. Id. The state argued that Bobby Joe’s statement that he had dreamed he had shot Bud was “competent to show ill will and malice against the deceased.” Id. The Supreme Court of North Carolina agreed with the defendant’s brief that stated, “‘We submit that even a highly specialized expert would have difficulty explaining the meaning of such a dream.’” Id. The court went on to note:

What a dream means, if anything, presents an occurrence filled with mystery. As to the meaning of a dream, we can only conjecture. The evidence as to the statement of the defendant that he dreamed that he shot Bud leaves the meaning of the dream in the realm of mere conjecture, surmise, and speculation, and one surmise may be as good as another. Nobody knows.

Id. at 724. The speculative nature of the dream, however, worked against the defendant. Id. Even if the evidence were incompetent, the court concluded, “it is, in our opinion, so speculative and uncertain as to have had no probative force on the minds of a jury and would not justify a new trial of this case.” Id.

Judge Ito’s application of the White court’s rationale—that dream evidence is so speculative it can have no probative force on the minds of the jurors—must have been his reason for admitting the dream talk evidence, although the record does not reflect his rationale. Ito merely stated, “The statement regarding the dreams about killing will be admitted.” Motion re Admissibility, supra note 49, at *13.
going to be asking them to do here.”76 Without explaining why, Judge Ito finally ruled that the references made to the polygraph test would be excluded, and that the “statement regarding the dreams about killing will be admitted.”77

To my mind, Judge Ito made a mistake.78 Simpson’s statement about

76. Motion re Admissibility, supra note 49, at *12.
77. Id. at *13.
78. In the press, there were many commentators who shared this view. The former Los Angeles County District Attorney, Ira Reiner, came right out and called Ito’s ruling “wrong”:

Before you can put on any evidence about the significance of our dreams, you have to deal with the question of do our dreams truly reveal our thoughts and feelings . . . Maybe they do, maybe they don’t. There is nothing approaching general agreement in the scientific community on this.

Jim Newton & Andrea Ford, Simpson Dreamed of Killing, Witness Says, L.A. TIMES, Feb. 2, 1995, at 1. A defense attorney, Gerald L. Chaleff, posited that if there is such a thing as reversible error, “it would seem like this is it. . . . Discussions of a defendant’s dreams or nightmares are prejudicial by their very nature because they suggest that merely dreaming about something means you’re going to do it. Common experience tells all of us that isn’t true.” Id. Professor Peter Arenella from the University of California at Los Angeles agreed that “Ito faltered badly in allowing the testimony.” McGory, supra note 69, at 3. However, Arenella went on to say, “It is clearly an error to admit this dream, for several reasons. This is a great issue for appeal, but in California, almost nothing is a reversible error. A California appellate court might still rule this is harmless.” Id. (internal quotations omitted).

Had a California court ruled the admission of dream evidence to be harmless error, it would not find itself alone. For example, in State v. Tyler, the Supreme Court of Kansas ruled that admitting a defendant’s dream narration via another witness was improper. 840 P.2d 413, 426 (Kan. 1992). In Tyler, the defendant was convicted of second-degree murder, aggravated assault of a law enforcement officer, and various drug-related charges. Id. at 419. John Tafoya, a defendant’s friend, had testified at a preliminary hearing that the defendant had “described a dream to him in which a Colorado detective almost captured him.” Id. at 425. After that dream narration, the defendant said that if the Colorado detective, Pat Crouch, came to get him, “‘[t]hey would have to shoot him and he would take someone out with him.’” Id. The Kansas Supreme Court held that it was an abuse of discretion to admit Tafoya’s testimony to prove the defendant’s state of mind. Id. at 426. The court reasoned: “Such evidence is too speculative to be reliable. Although Tyler’s statement to Tafoya was not a part of the actual dream, it was so closely related to the dream that it also lacks probative value.” Id. The court concluded, however, that the error was harmless. Id.

In a Michigan appellate court case, People v. Allen, No. 212699, 2000 WL 33418839 (Mich. Ct. App. June 9, 2000), the defendant was convicted of a first-degree murder for shooting his codefendant’s girlfriend who was preparing to inform the police about her boyfriend’s drug dealing. Id. at *1. After the shooting occurred, the defendant spent the night with friends. Id. He awakened his friend Riddle crying. “‘She was after me, and that she is all bloody.’” Id. Later Riddle told someone else that she had “‘heard the dream he had.’” Id. Citing Tyler, the court agreed that the evidence of defendant’s dream, and his statement following that dream even though he was awake when he made the statement, lacks probative value. The statement was made immediately upon defendant awakening and was clearly induced by the dream. “Such evidence is too speculative to be reliable.”

Id. at *2 (quoting Tyler, 840 P.2d at 426). However, also following the rationale in Tyler, the court held that the error was harmless. Id. It could not “conclude that Riddle’s testimony concerning defendant’s statement about his dream had such an effect as to undermine the reliability of the
having dreamed of killing Nicole should not have been admitted. That statement directly related to Simpson’s fear about taking the lie detector test. He was afraid that he might fail the test. He knew that polygraph tests measure changes in the subject’s body that are beyond the subject’s control: damp palms, beating heart, shallow breaths.\textsuperscript{79} When it came time to answer questions about killing Nicole, he might balk because of the dreams he remembered about killing her. Just as a dog might twitch involuntarily in his sleep over chasing a rabbit, its body being fooled into thinking the dream was really happening, Simpson was worried that the opposite would happen. While awake, his body might twitch involuntarily when asked about killing Nicole, being fooled into thinking the dream had really happened. It was not an irrational fear at all. Those involuntary twitches might look a lot like a confession.\textsuperscript{80}

Simpson’s thought had two parts: (1) I am afraid to take the polygraph test because (2) I have had a lot of dreams about killing Nicole. Ito, and all of the attorneys involved, knew that the first part of the thought was inadmissible because of the reference to the polygraph test. The second part of the thought, however, explained the reason for the first part. Indeed, had the occasion for the first part never arisen, the second part might never have been said. To allow the statement embodying the second part of the thought to reach the jury without admitting the first part is out-of-context, misleading, and highly verdict.” \textit{Id.} The dream talk was also admitted in the trial of his co-defendant, relying on \textit{Tyler.} See \textit{People v. Shaw, No. 211825, 2000 WL 33533976, at *3 (Mich. Ct. App. Feb. 29, 2000).}

\textsuperscript{79} “The polygraph is a device that measures physiological reactions of humans in an attempt to determine the veracity of statements they make.” \textit{Timothy B. Henseler, Comment, A Critical Look at the Admissibility of Polygraph Evidence in the Wake of Daubert: The Lie Detector Fails the Test, 46 CATH. U. L. REV. 1247, 1251 (1997).} The polygraph instrument measures blood volume, heart rate, respiratory activity, and galvanic skin resistance (palm sweating). The examiner attaches the instruments to the subject[,] any changes in the physiological measurements are transmitted to pens which record them on moving chart paper.

The general hypothesis upon which the theory of polygraphy rests is the notion that when a person lies, the human aversion to lying causes a physiological response which in turn causes an involuntary physiological reaction. A lie is believed to precipitate an alteration in the rate and pattern of breathing, blood pressure, rate and volume of blood flow, and the moisture on the skin. Truthful responses do not precipitate these same reactions. \textit{United States v. Cordoba, 194 F.3d 1053, 1057 (9th Cir. 1999).}

\textsuperscript{80} It is also possible that embarrassment about having had such a dream might cause Simpson’s body to respond involuntarily in response to questions about murdering his wife. However, given the context of the polygraph examination, as an investigatory tool for the murder of his wife, it is more likely that Simpson was fearful that his body’s involuntary responses were going to be interpreted as evidence that he had in fact killed her.
Wild Dreamers

prejudicial. Lacking proper context and therefore failing to explain why he was afraid to take the polygraph test, the statement about dreaming of killing Nicole looks like an unprompted party admission. And it does not take much imagination to see how easily a juror might make the mental leap from an out-of-context admission to an out-and-out confession.

To my mind, both parts of Simpson’s thought should have been excluded: the first part because evidence about polygraph tests is inherently unreliable and must be excluded, and the second part because the first part had to be excluded. The first part gave meaning and context to the second part. This contextual argument alone should have persuaded Judge Ito to exclude the evidence. But there is another more compelling reason: the jurors might have ascribed alternate meanings to O.J. Simpson’s dreams. From those alternate meanings, they may have unfairly drawn an inference of guilt.

C. A Freudian Interlude: Simpson’s Fatal Obsession

It is time for a Freudian interlude. The prosecutor’s main argument in favor of admitting Simpson’s dream talk was that it evidenced a “fatal obsession” for Nicole. This theory of relevance obliquely refers to

81. A sense of unfairness about introducing bits and pieces of an out-of-context statement is embodied in Federal Rule of Evidence 106: “When a writing or recorded statement or part thereof is introduced by a party, an adverse party may require the introduction at that time of any other part or any other writing or recorded statement which ought in fairness to be considered contemporaneously with it.” Fed. R. Evid. 106. Federal Rule of Evidence 106 only applies to writings or recorded statements. Id. For oral conversations, the opposing party must develop the remainder on cross-examination, or as part of his own case. See Beech Aircraft Corp. v. Rainey, 488 U.S. 153, 171–72 (1988). The problem in the Simpson trial was that the defense did not, and could not because of the evidence rules, introduce the statement about the polygraph that would have given context and explanatory meaning to the statement about the dream.

82. For a fascinating discussion of the O.J. Simpson case and its parallels to the Salem witchcraft trials, see Peter Charles Hoffer, Invisible Worlds and Criminal Trials: The Cases of John Proctor and O.J. Simpson, 41 AM. J. LEGAL HIST. 287 (1997). Hoffer points out that while the rules of evidence and modern trials bow to “the bright light of science and the expert witness,” id. at 290, there are invisible worlds still influencing jurors. For middle-class whites, that invisible world includes the fear of crime and the identification of criminals as dangerous and black; their invisible world created a conspiracy theory based on “black malevolence.” Id. at 313–14. For black jurors, however, it was easy to convince them of the conspiracy of the police to nail Simpson, since for them, “South Central is a world in which the police always and routinely bend and break the rules— much like the world where the Devil and his minions do their mischief.” Id. at 292. Thus, both whites and blacks created their own conspiracy theories that were generated by their respective invisible worlds.

83. Motion re Admissibility, supra note 49, at *3.
Sigmund Freud’s theory that a “dream is the fulfilment of a wish.” For Freud, dreams had meaning:

[They are] not to be likened to the unregulated sounds that rise from a musical instrument struck by the blow of some external force instead of by a player’s hand; they are not meaningless, they are not absurd; they do not imply that one portion of our store of ideas is asleep while another portion is beginning to wake. On the contrary, they are psychical phenomena of complete validity—fulfillments of wishes; they can be inserted into the chain of intelligible waking mental acts; they are constructed by a highly complicated activity of the mind.

According to Freud, the human organism is plagued by a host of primitive needs and desires that are threatening to the personality. Our id is instinctual, biologically driven—we want, we desire, we need, we have to have. Our superego is the internalized parent who commands socially appropriate behavior and seeks to control our hidden impulses. To the ego is left the task of reconciling the two forces of the id and the superego, as well as responding to the demands of outside realities. When one’s ego reaches a level of maturity, the demands of the id and the superego are not denied, but are dealt with in a fashion that is no longer destructive. Getting the ego into a position of healthy defense against the tyranny of the id and the superego and against the demands

84. FREUD, supra note 47, at 154.
85. Id. at 155.
86. These primitive needs and desires are often distressing to the dreamer, and that attitude of repugnance towards them requires the dreamer to distort them. As Freud wrote,

Everyone has wishes that he would prefer not to disclose to other people, and wishes that he will not admit even to himself. On the other hand, we are justified in linking the unpleasurable character of all these dreams with the fact of dream-distortion. And we are justified in concluding that these dreams are distorted and the wish-fulfillment contained in them disguised to the point of being unrecognizable precisely owing to the repugnance felt for the topic of the dream or for the wish derived from it and to an intention to repress them. The distortion in the dream is thus shown in fact to be an act of the censorship.

Id. at 193.
88. Id. at 642–43.
89. Id. at 656.
of external reality is presumably what we mean by “growing up.”  It is also the goal of psychoanalytic therapy.

In 1900, Freud published *The Interpretation of Dreams*, and his treatment of dreams illuminated his theory of the unconscious. Freud believed that the unconscious is hidden from and unobservable by the conscious mind. Much of what exists in the unconscious is there through the process of repression. The mind pushes beneath the surface thoughts that are too painful or dangerous to contemplate, such as incestuous sexual desires or violence to others. Drives that cannot find an outlet are then often rechanneled into socially acceptable behavior through a process known as sublimation. Similarly, thoughts that are

91. *Id.*

92. *Id.* According to Freud, in interpreting the dream, the dreamer reports the manifest content of the dream to the analyst, and then proceeds to associate freely to various parts of the dream. Richard M. Jones, *Freudian and Post-Freudian Theories of Dreams*, in *HANDBOOK OF DREAMS*, supra note 14, at 271, 276–77. The analyst then seeks to draw inferences about how the unconscious wish linked itself up to the residue of images from the previous day in order to discover the latent content of the dream. *Id.* at 277. As one scholar put it, “The sole aim of the interpretive process is to reformulate the manifest content of the dream back into the latent thoughts which, it is assumed, provoked the dream in the first place.” *Id.*

93. *FREUD*, supra note 47. While Freud squarely placed the dream within the scientific domain, his work was part of a trend in the nineteenth century psychological literature to show an interest in dreams. Webb, *supra* note 41, at 180. One scholar, Hendrika Vande Kemp, looked at American and British psychological periodicals, both popular and professional, between 1860 and 1910, and discovered that there was a steady increase in articles about dreams from 1860 to 1870. *Id.* Thereafter, interest in dreams declined in popular periodicals but increased in professional journals. *Id.* By 1910, dream literature was almost exclusively found in professional journals. *Id.*

94. Freud’s *The Interpretation of Dreams* was published in November 1899, although he had been working on the book since about 1892 when he began to record his analyses of dreams. See Peter Gay, *Introduction to SIGMUND FREUD, The Interpretation of Dreams*, in *THE FREUD READER*, supra note 87, at 129, 129. The most celebrated of the dreams was the so-called “Irma dream,” which was dreamed in 1895. *Id.* After his interpretation of this famous dream about the incurable Irma, a personal friend and patient of Freud, Freud wrote:

The dream acquitted me of the responsibility for Irma’s condition by showing that it was due to other factors—it produced a whole series of reasons. The dream represented a particular state of affairs as I should have wished it to be. *Thus its content was the fulfillment of a wish and its motive was a wish.*

*FREUD*, supra note 47, at 151.

95. *FREUD*, supra note 47, at 651 (“The unconscious is the true psychical reality; in its innermost nature it is as much unknown to us as the reality of the external world, and it is as incompletely presented by the data of consciousness as is the external world by the communications of our sense organs.”).

96. *See id.* at 268–70.


98. *See SIGMUND FREUD, Leonardo Da Vinci and a Memory of His Childhood*, in *THE FREUD READER*, supra note 87, at 443, 452.
too painful or dangerous to contemplate often show up in our dreams. According to Freud, the human organism is plagued by repressed infantile wishes that create “intolerable states of psychological tension in waking life.” These unconscious wishes or motives, expressed in symbolic form, are hidden in the manifest content of the dream—what the dream was apparently about.

Freud believed that two psychical forces operate to give dreams shape. The first force “constructs the wish which is expressed by the dream.” However, while this unconscious wish is too strong to be denied, it is too threatening to be directly expressed. Thus, the second force “exercises a censorship upon this dream-wish and, by the use of that censorship, forcibly brings about a distortion in the expression of the wish.” The second agency censors out “distressing content,” allowing “nothing to pass without exercising its rights and making such
Wild Dreamers

modifications as it thinks fit in the thought which is seeking admission

to consciousness." 106 The function of this second agency is “of a
defensive and not of a creative kind.” 107 The wish emerges in symbolic
form, a tangle of internalized prohibitions and repressions.

Furthermore, the sexual nature of our most basic wishes results in
dreams full of sexual symbolism that often reflect hidden incestuous
desires. 108 While Freud denied that all dreams require a sexual
interpretation, he also asserted that a “great many other dreams,
however, which show no sign of being erotic in their manifest content
are revealed by the work of interpretation in analysis as sexual wish-
fulfillments.” 109 Thus, most dreams are about inexpressible and
threatening sexual desires. 110 Since these dreams are too awful to face,

106. Id. at 177–79.

107. Id. at 179.

108. One of the processes that takes place in dream formation Freud calls “condensation.”
FREUD, supra note 101, at 151. Various constituents of the dream become compressed, and for any
single element in a dream, there are a number of different associative threads that “branch out in
two or more directions; every situation in a dream seems to be put together out of two or more
impressions or experiences.” Id. The material in dream thought is “packed together for the purpose
of constructing a dream-situation,” and there is a common element among them. Id. at 152. Thus,
in analysing a dream, if an uncertainty can be resolved into an ‘either—or’, we must replace it for
purposes of interpretation by an ‘and’, and take each of the apparent alternatives as an independent
starting-point for a series of associations.” Id.

109. Id. at 169–70.

110. Freud believed that the adult retained the infantile forms of sexual life, and that repressed
infantile sexual wishes were the most “frequent and strongest motive-forces for the construction of
dreams.” Id. at 170. The sexual ideas in the dream are not represented as such, “but must be
replaced in the content of the dream by hints, allusions and similar forms of indirect representation.”
Id. There are stock “dream-symbols,” which

serve to represent persons, parts of the body and activities invested with erotic interest; in
particular, the genitals are presented by a number of often very surprising symbols, and the

greatest variety of objects are employed to denote them symbolically. Sharp weapons, long and
stiff objects, such as tree-trunks and sticks, stand for the male genital; while cupboards, boxes,
carriages or ovens may represent the uterus.

Id. at 171.

We do not know the form of murder that Simpson dreamed about, but perhaps his dreams were of
stabbing Nicole with a sharp object—the phallic symbolism of the knife is impossible to miss. Most
of Freud’s discussion about murder, however, concerns the son’s murder of his father. Freud even
took the position that the commandment “thou shalt not kill” first referred to the most murderous of
acts, the killing of the primitive father. See SIGMUND FREUD, The Future of an Illusion, in THE
FREUD READER, supra note 87, at 685, 712 (“Under totemism, this commandment was restricted to
the father-substitute, but it was later extended to other people, though even to-day it is not
universally obeyed.”).
we instead dream about something else that is symbolically a representation of that unexpressed desire.111

Who knows what Freud would have said about Simpson’s dreams? From my cursory reading of Freud, it seems that if we could resurrect the gentleman and put him on the stand, his interpretation might have had something to do with Simpson’s unresolved and threatening fear of being in love with his mother and his anxiety about being castrated.112 But Freud is dead and difficult to subpoena. In any case, we are not now concerned with Freud’s interpretation—we are concerned with the interpretations of those diverse members of the jury.

The impact of Freud’s legacy is undeniable. His insights about the unconscious have by now become mainstays in popular culture. Freudian jargon and concepts have percolated down to almost anyone who reads books or magazines, or who watches film or television, even if just the commercials. We can expect that some members of Simpson’s jury knew about and believed in the power of the unconscious to motivate conscious action, and that dreams were a manifestation of that unconscious. Moreover, some jury members likely had misconceptions about Freud’s theories. Their inferences about Simpson’s “wish-fulfillment” were likely far more simplistic, and far more dangerous to Simpson’s defense, than the more convoluted exercises in dream symbolism that Freud himself would employ. Some jurors may well

111. Freud believed that dreams were derived from the events of the preceding day, or past few days. FREUD, supra note 101, at 155 (“[W]e find that every dream without any possible exception goes back to an impression of the past few days, or, it is probably more correct to say, of the day immediately preceding the dream, of the ‘dream day.’”). If a connection between the content of the dream and any impression of the previous day is made, “that impression is so trivial, insignificant and unmemorable, that it is only with difficulty that we ourselves can recall it.” Id. Since the dream seems to be concerned with “the most indifferent trivialities,” it therefore appears to be “unworthy of our interest if we were awake. A good deal of the contempt in which dreams are held is due to the preference thus shown in their content for what is indifferent and trivial.” Id. However, “analysis uncovers the numerous associative paths connecting these trivialities with things that are of the highest psychical importance in the dreamer’s estimation.” Id. at 156. Through the distorting process of displacement, “the psychical intensity passes over from the thoughts and ideas to which it properly belongs on to others which in our judgment have no claim to any such emphasis.” Id. at 154–55.

112. According to Freud, adolescent boys are anxious about being in love with their mothers, and fearful of being punished by castration. Sigmund Freud, Anxiety and Instinctual Life, in The Freud Reader, supra note 87, at 773, 777 (“But we have not made any mention at all so far of what the real danger is that the child is afraid of as a result of being in love with his mother. The danger is the punishment of being castrated, of losing his genital organ.”). Freud insists this fear of castration cannot be dismissed lightly: “He has some ground for this, for people threaten him often enough with cutting off his penis during the phallic phase, at the time of his early masturbation.” Id. at 777–78.
have reasoned: Simpson dreamed of killing his wife; therefore, Simpson wanted to kill his wife.113 In that same line, the prosecutor argued, sometimes wishes become so powerful that they become a “fatal obsession.”114 Id over superego. Bingo: Nicole is dead. O.J. dreamed of doing it; therefore, he wanted to do it; therefore, he did it.115

The Freudian interlude concludes. But there are other, wilder theories about the meaning of O.J. Simpson’s dreams that worry me even more than watered down, ill-formulated Freudian wish-fulfillment.

D. Other Wilder, More Worrisome Theories About the Meaning of O.J. Simpson’s Dreams

These theories are less likely to be articulated in a court of law where reason and an empirical episteme reign supreme. They are based on the concept of the dream as a form of hearsay and implicate a premodern or medieval mindset.116 Hearsay is defined as an out-of-court statement

113. Here is how some jurors might have reasoned: Simpson’s dreams of murdering Nicole were indicative of Simpson’s murdering state of mind. He had such hostility for his wife that he wished her a violent death; furthermore, he wished to be the instrumentality of that violent death. His dreams manifested these latent desires. While they may not have explained the motive for her violent death—why he came to hate her so much—the dreams surely evidenced an inchoate mens rea. And they surely suggest the identity of her murderer: Who in the world could hate Nicole Simpson so much that he wished to destroy her in such a violent way? The dreamer of hateful dreams—the defendant, O.J. Simpson.

114. See supra note 65 and accompanying text.

115. O.J. Simpson would not have fared much better had the members of the jury applied an alternative theory about dreams: Alfred Adler’s. Adler’s theory about dreams and their function differed from Freud’s. Adler viewed the dream as integral to the thought process, and reflective of the dreamer’s view of life, combined with his goals. Leo Gold, Adler’s Theory of Dreams: An Holistic Approach to Interpretation, in HANDBOOK OF DREAMS, supra note 14, at 319, 321. The dream had a “forward aim, that it ‘puts an edge’ on the dreamer for the solution of a problem in his own particular way.” Id. (quoting ALFRED ADLER, SOCIAL INTEREST: CHALLENGE TO MANKIND 259 (1938)). The dream represents unfinished business and carries into sleep all the impression and events of the day. Id. There is a continuity between our waking and sleeping thoughts, and the individual is always geared towards coming up with possible solutions to the mundane demands of life, as well as the development of long-term goals. Id. at 326. The dream, like all thought, is part of the process of achieving the future. Adler stated, “In dreams we reproduce the pictures which will arouse the feelings and emotions which we need for our purposes, that is, for solving the problems confronting us at the time of the dream, in accordance with the particular life style which is ours.” Id. at 325–27 (quoting H.L. ANSBACHER & R.R. ANSBACHER, THE INDIVIDUAL PSYCHOLOGY OF ALFRED ADLER 361 (1956)) (internal quotations omitted). Hence, Simpson’s dreams about killing his wife would have been considered part of the process of achieving his future—presumably a future that did not include Nicole.

116. See supra notes 35–44 and accompanying text.
offered to prove the truth of the matter asserted.\textsuperscript{117} Even though it would never be said out loud in court, I suspect that the truth of the matter sought to be proved by the prosecution in Simpson’s narration of the dream was—depending upon how one defines the relevant time frame—Simpson has killed Nicole, or Simpson will kill Nicole.

The analysis gets complicated. It is possible that Shipp’s testimony about Simpson’s narration of the dream could be considered multiple hearsay. The first level of hearsay, and the only level recognized by the court, is obvious: Simpson’s out-of-court narration of the dream to Shipp.\textsuperscript{118} But the second level of hearsay is subtler: Simpson’s dream itself could be regarded as an out-of-court statement.

Who is the declarant of this second level of hearsay—the dream itself? Under a premodern or medieval mindset, the declarant of the dream itself may be God, or some other invisible, transcendent entity, a demon or an alien perhaps. This unknown entity was the author of the dream, or in evidentiary terms, someone who had made a statement in the form of a dream that Simpson experienced in his sleep. Under this analysis, we have Ron Shipp who would testify that he personally observed O.J. Simpson’s narration of the dream. Shipp would say that Simpson told Shipp, “I was an eye-witness to this dream, although I was asleep and my eyes were closed. I personally observed the making of the second declarant’s statement, even though I don’t know who the second declarant was.” Thus, there were two levels of hearsay: the first was the narration by Simpson of the dream, the second was the statement by an unknown, possibly supernatural, declarant that constituted the dream itself.

The danger of this analysis is in the second level of hearsay—the dream authored by an unknown entity. That statement seemed to say: O.J. Simpson has killed or will kill his wife—which is, of course, the very truth that the prosecution is seeking to prove. The law of evidence is settled about how to deal with traditional instances of multiple hearsay.\textsuperscript{119} It tackles the issue every time a police officer makes a written

\textsuperscript{117} The definition of hearsay is found in Federal Rule of Evidence 801. It first defines the term "statement" as: “1) an oral or written assertion or 2) non-verbal conduct of a person, if it is intended by the person as an assertion.” \textit{Fed. R. Evid.} 801(a). Then “hearsay” is defined as: “a statement, other than one made by the declarant while testifying at trial or hearing, offered in evidence to prove the truth of the matter asserted.” \textit{Id.} 801(c).

\textsuperscript{118} See \textit{also supra} note 63.

\textsuperscript{119} Federal Rule of Evidence 805 provides: “Hearsay included within hearsay is not excluded under the hearsay rule if each part of the combined statements conforms with an exception to the hearsay rule provided in these rules.” \textit{Fed. R. Evid.} 805.
report of a statement that was made to him at the scene of an incident. Let us imagine a civil lawsuit arising out of a car accident. Person $X$ makes Statement #1 to Police Officer $Y$ about what he observed at the time of the accident. Then Police Officer $Y$ writes down Statement #2 in his incident report that reads: “Person $X$ says, ‘Statement #1.’” At trial, neither declarant is present to testify; hence, both statements are hearsay. Each level of hearsay must qualify for an exception to the hearsay rule in order for the document containing Statement #1 to be admissible. \(^\text{120}\) We might get Police Officer $Y$’s Statement #2 into evidence under some version of a public records exception, \(^\text{121}\) and then we have to deal with Statement #1: was it perhaps some form of spontaneous declaration, \(^\text{122}\) a party admission, \(^\text{123}\) or some other kind of recognized statement with sufficient reliability to let the jury hear it?

But when the second declarant is an unknown, possibly supernatural entity, as in our premodern scenario, how do the Rules of Evidence govern? We can leave that question unanswered—conceiving of the dream itself as a level of hearsay is never going to happen. Lawyers are running the trial on all fronts, bench and bar. Lawyers are consummate rationalists, thoroughly indoctrinated in the dominant episteme. It will never occur to most of them that the dream could be considered a statement—except on some uncomfortable, unconscious level.

But here is the rub: some members of the jury may not swear fealty to the dominant episteme. They may also not be the mavens of rational thought that the law of evidence assumes them to be. They may know credible wild dreamers, or be wild dreamers themselves. To them, the defendant’s narration of dreams that he killed his wife, over and over again, may be heard as double hearsay—under a premodern theory, with the dream constituting a second declarant’s statement that O.J. Simpson either has or will soon kill his wife. They may believe that God, or some other invisible transcendent being, was trying to transmit information, and such a belief would be incredibly harmful to the defense.

\(^{120}\) See id. This rule about multiple hearsay applies in California as well. See CAL. EVID. CODE § 1201 (West 1995).

\(^{121}\) Federal Rule of Evidence 803(8), for example, provides an exception to the hearsay rule for certain records and reports prepared and maintained by public offices and agencies. FED. R. EVID. 803(8). Note that Federal Rule of Evidence 803(8)(B) imposes a restriction in criminal cases upon the use of records of reports from police officers and other law enforcement personnel containing matters observed pursuant to official duty. Id.

\(^{122}\) See id. 803(2) (providing excited utterance exception to hearsay rule).

\(^{123}\) See id. 801(d)(2) (defining party admissions as not hearsay). I am grateful to Professor Gary Shaw for his feedback on this double hearsay analysis.
This harm suggests that even if the court cannot bring itself to consider a dream a statement, the fact that jurors might well do so should preclude its admissibility under Federal Rule of Evidence 403.\textsuperscript{124} There are times when an item of evidence may be probative, at least logically, of some material fact, but it should be excluded because it poses a substantial risk of unfair prejudice if heard by the jury.\textsuperscript{125} This is a catch-all category in which to dump arguably relevant evidence that might evoke horror or repulsion, or prod the jurors into unduly punishing the defendant because they cannot get some image or damaging idea out of their minds.\textsuperscript{126}

Under Rule 403, Judge Ito should have excluded any reference to Simpson’s repeated dreams of murdering his wife. All sorts of illogical propositions may have flashed through the minds of the jurors, silently in their inner minds, or publicly within the walls of the deliberation room. But the jurors’ inner minds may have been secretly or unconsciously wedded to an alternative episteme, one more medieval than modern, or to something else entirely. By hearing about the defendant’s dreams in the courtroom, the seeds of thought were planted. Whether or not the full-grown plants make sense to the rational empiricist does not matter. They may have grown like powerful weeds in the untamed garden of the inner mind—tough, reedy, impossible to extirpate.

O.J. Simpson was a wild dreamer extraordinaire, a famous wild dreamer who became entangled with the law, but who was ultimately

\textsuperscript{124} Id. 403 (“Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.”). California has its analog of Federal Rule of Evidence 403, which provides: “The court in its discretion may exclude evidence if its probative value is substantially outweighed by the probability that its admission will (a) necessitate undue consumption of time or (b) create substantial danger of undue prejudice, or confusing the issues, or of misleading the jury.” CAL. EVID. CODE § 352.

\textsuperscript{125} The Advisory Committee Notes to the Federal Rules define “unfair prejudice” as “an undue tendency to suggest decision on an improper basis, commonly, though not necessarily, an emotional one.” FED. R. EVID. 403 advisory committee’s note.

\textsuperscript{126} Some examples of cases involving evidence deemed inadmissible due to prejudice are Wallace v. Mulholland, 957 F.2d 333, 336 (7th Cir. 1992) (refusing to admit evidence of a defendant’s schizophrenia in a civil rights case involving police injury to a mental patient), United States v. Masters, 924 F.2d 1362, 1368 (7th Cir. 1991) (refusing to allow evidence of alleged transvestism of a third party in a murder trial), United States v. Eason, 920 F.2d 731, 738 (11th Cir. 1990) (refusing to allow as evidence the fact that the defendant’s father was convicted of a similar offense), and In re Bendectin, 624 F. Supp. 1212, 1224 (S.D. Ohio 1985) (refusing to allow the presence of a deformed child in the courtroom).
Wild Dreamers

acquitted. But not all wild dreamers are so lucky. I will now tell the stories of three wild dreamers from family court where nightmares abound, lesser mortals who narrated their dreams to others and came to rue the day. One man dreamed of his wife’s car exploding; the second merely dreamed of days to come. The third wild dreamer was a six-year-old child who dreamed of her father killing her—and then he did. It is through the prism of her tragic situation that we will look at Carl Jung’s attitude towards dreams, and at the dreaming culture of the seventeenth century Iroquois.

II. THREE WILD DREAMERS FROM FAMILY COURT, A JUNGIAN INTERLUDE, AND THE DREAMING CULTURE OF THE IROQUOIS

A. Wild Dreamer #1 in Family Court Who Dreams of His Wife Exploding

Susan V.H. v. Robert W.H. involved a custody dispute before a Delaware family court, a battle over who would raise the shattered child of a shattered marriage. The father, Robert H., appears to have been an abusive alcoholic. In January 1997, the court issued an Order of Protection from Abuse, providing that the father shall not “threaten . . . harass or commit any other act of abuse” against the mother, and shall stay “100 yards away” from her, and shall not attempt to contact her in any way. A second order, issued the following March, provided that the father shall not consume any alcohol during the twenty-four hours preceding any visitation with their son, Austin, who was two years old at the time. Robert H. failed to comply with both

128. That it was family court came as no surprise to me. In my brief foray into practice, I did an internship with the Murder Task Force of the Chicago Public Defender, and then several years of family law. Between the two, I used to say: give me criminal law over matrimonial—any day. Criminal practice was a lot cleaner, and for the most part, my clients, the indigent men and women accused of murder, were a lot saner, and a lot nicer, than my clients in family court.
130. See generally id.
131. Id. at *1.
132. Id. (alteration in original) (internal quotations omitted).
133. Id.
orders. Following one of the hearings, he came up to the mother, Susan H., and told her that if his visitation privileges stopped, he would leave the state and stop paying child support. Susan H. also testified that Robert H. had told her that he “think[s] of the day . . . [that her] car will explode.”

Robert H. testified himself about the conversation. Susan H. had parked her car beside his, and waited for him following their child support hearing. The two talked for “approximately one hour, during which he told [Susan] that he ‘dreams’ of her car exploding with her in it.” Robert H. elaborated that this dream is what “keeps . . . [him] clinging to life.” The court held that the second statement was “abuse” under Delaware law, and ordered Robert H. to submit to a psychological evaluation. In addition, because the judge found the father’s “dreams’ so chilling,” the matter was referred to the Office of the Attorney General for “possible criminal prosecution.”

We do not know how Robert H. was using the term “dream.” Were these images of Susan H.’s exploding car just dreams Robert H. had in sleep—plain old, garden variety Freudian death wishes? Or perhaps they were daydreams, fantasies he engaged in while drifting through what William James calls the “dreamy states.” The referral to the Office of the Attorney General suggests that the judge regarded these images, and the narration of them, as potential threats to Susan H.’s life. But does it seem appropriate to refer the matter to the prosecutorial arm of the state for investigation on the basis of Robert H.’s set of purported “chilling dreams”?

There is great danger in taking that step. If we were held criminally liable for our murderous dreams, then many of us would end up in

134. Id. at *2.
135. Id.
136. Id. (alteration in original). He admitted that “he told Mother that he dreams of the day when her car will explode with her in it.” Id.
137. Id.
138. Id.
139. Id.
140. Id.
143. Id.
144. JAMES, supra note 5, at 295.
prison. How many Presidents of the Rotary Club have woken in a nocturnal sweat from dreaming of the violent death of his beloved wife who lies beside him—the same beloved wife who will get up in the morning and make him blueberry pancakes? But clearly Robert H. is not the president of any civic organization, and from reading the bare facts of the case as reported in the appellate decision, I would have no bone to pick with any judge who ordered Robert H. to keep a respectful and safe distance from his wife and child, or who ordered a psychological examination. I could probably never be talked into referring the matter to the prosecutor’s office, but someone might give it a try, under a different analysis.

It is not the judicial action with which I quarrel, but its justification. Instead of basing his order on Robert H.’s “chilling dreams,” the judge should have based it on Robert H.’s attitude towards those dreams. Even applying Freud’s wish theory, Robert H.’s dreams of Susan H.’s car exploding do not necessarily tell us anything about how he feels about his wife. Without the dreamer being an agent in causing her destruction, the dream could represent nothing but fear of loss. But telling us that those dreams are what keep him “clinging to life” is another matter altogether. How he feels about her is revealed. And in the context of a past history of alcoholism, and presumably violent behavior towards his wife, I agree with the judge—his attitude towards the dreams of her destruction is chilling. Maybe I could even issue an order to have Robert H. investigated for that statement as a not-so-thinly-veiled threat.

By fine-tuning the basis for justifying the judge’s order, we avoid the muddy waters of conflicting epistemes and murky metaphysics. No one has to question what the dream might mean, only what the dreamer’s attitude towards the dream might mean. Not only is this analysis fairer to folks like Robert H., it also keeps the President of the Rotary Club out of prison for dreaming about the violent death of his wife. He can dream without worry or censorship. He might even wish that she were dead, even if her blueberry pancakes are the best. But he should beware. Once he starts to make noises about clinging to those dreams for his dear life, some of us may begin to worry about his dear wife’s life, and might condone invoking the power of the state.

And what if the wild dreamer’s dreams have no violent content, but he is capable of telling the future from his dreams? Let me tell you what happened to Terry G.
B. Wild Dreamer #2 in Family Court Who Only Dreams of Tomorrow

It was a custody case, and as is often true in these cases, neither parent in In re Interest of Amber G. v. Terry G. had much to offer. Following the Nebraska divorce, Amber, Jessica, Adam, and Brittany went to live with their mother, Dolly G.; their father, Terry G., then went to live in Missouri for awhile. Following reports of sexual abuse and neglect, the children were placed in foster care, under the temporary custody of the Department of Social Services (DSS). For three years, DSS tried to reunite the family. Visitations by both parents occurred on a regular basis, but the DSS reports recommended counseling for both parents, indicating the “past abuse and chaotic or dysfunctional life styles by both parents,” and noting that “placement [with the father] had been questioned as marginal due to past history,” and that while both parents’ lifestyles “must change to provide safety and stability . . . [t]his is questionable due to the ongoing instability both have or have had.”

There was never much of an issue regarding the mother: the children were not going back with her. Rather, the choice seemed to be whether the children would stay in foster care under a permanent guardianship, or would live with their father. Terry G. was no treat either. He was threatening to the foster parents and the social worker, used foul and abusive language with both, frightened his children, allowed them to hit each other, stated that he hated his son, let them wear feces-stained underwear and other unwashed clothing, never washed their hair, and let them take their own medication without supervision. His philosophy for managing his children was laissez-faire—and hostile: “When kids

145. 554 N.W.2d 142 (Neb. 1996). The father had appealed the order of the county court recommending permanent guardianship of his four minor children. Id. at 145.
146. See generally id.
147. Id. at 145.
148. Id. at 145–46.
149. Id. at 146.
150. Id.
151. Id. at 145–46.
152. Id. at 146 (“[D]ad said he could knock [the social worker] clear across the street. That he wanted to push his fist down [the foster mother’s] throat. Lot of swearing, such as, and pardon my language[,] asshole, shit used frequently. Threaten to sue the [foster family], accusing them of taping conversations. That he was going to start trouble in Clearwater.”).
153. Id.
Ain’t fightin’ or running away, I don’t give a shit what they do.\textsuperscript{154} Anyone reading the DSS report would have agreed with the social worker: Terry G. was unfit for parenting.

But here is the part of the case that bothered me. Terry G. testified that he had dreams in which he could predict the future:

I go to sleep at night. About five minutes, if I’m asleep I see things. And if—I don’t want to see them, so I wake up. But they just keep coming back. Three nights in a row. And when they do, then I know it’s going to happen. That’s all I can say.\textsuperscript{155}

Later in the case, we see that the father’s revelations about his prophetic dreams influenced the social worker in his decision not to recommend that Terry G. have custody: “In his meetings with the father, the clinical social worker expressed significant concerns about the father’s fitness to have custody due to his claims that he can predict future events through his dreams and his lack of awareness as to the significant needs of his children.”\textsuperscript{156}

No one would argue with the social worker’s considerable concerns about Terry G.’s lack of awareness as to the significant needs of his children. Those dreadful details—abusive language, the threats, the unsupervised medication, the unwashed hair, the dirty underwear—all warrant a finding that these kids would be better off in foster care. But what about those prophetic dreams of Terry G.? What relevance do those dreams have to his fitness as a father? I find myself looking at my children, and shuddering. I make a vow to keep my dreaming life to myself—a vow I promptly break. Dreams of things to come have on occasion happened to me. I know other kind and responsible parents who have experienced dreams that came true, or who have confessed to having a \textit{déjà vu}. We should all vow to keep our dreaming life to ourselves. We are all in jeopardy.

Here was an empiricist social worker so hostile to an alternative episteme that he deemed dream talk of this sort as evidence of parental unfitness. This social worker ranked his concern about Terry G.’s prophetic dreaming not only in the same sentence as his concern over Terry G.’s lack of awareness of his children’s needs, but in the breath before it. Then, to make matters worse, the appellate court chose to include those words of concern over the prophetic dreams in its decision.

\textsuperscript{154} Id. at 147.
\textsuperscript{155} Id.
\textsuperscript{156} Id. at 149–50.
affirming the denial of custody.\textsuperscript{157} To include the empiricist social worker’s assessment of unfitness due to wild dreaming is to ratify it, and to set precedent for other prophetic dreamers—who may be perfectly good parents—that they had better keep their mouths shut if they want to keep their children.

Don’t get me wrong: those children do not belong with Terry G. He was an unfit parent. There were plenty of facts about his conduct in the waking world to establish that. Where he travels in his sleep, however, and what information he garners in his dreaming state, from whatever source, extrasensory or divine, I am not willing to allow a court to consider. I fear he will be punished for having an alternative episteme, for being a wild dreamer. I fear they will take his children away. I fear they will take my children away.

But now let us turn to our third wild dreamer from family court. Her story is the saddest of all.

C. Wild Dreamer #3 in Family Court Who Dreams of Her Own Murder (and Then Is Killed)

This wild dreamer started out in family court, and then became the victim in her father’s murder trial.\textsuperscript{158} Only six years old, Ayla was a dreaming girl.\textsuperscript{159} She lived with her mother, Joann Daigle-Moylan.\textsuperscript{160} Her father, Diego Vas, who was not married to her mother, had supervised parenting time with Ayla at the Child Protection Council (CPC) in Danielson, Connecticut.\textsuperscript{161} In November 1992, Joann brought Ayla to the CPC office for parenting time, to be supervised by a CPC social worker, Joyce Lannan.\textsuperscript{162} During the session, Vas became visibly

\textsuperscript{157}. Id. Some scholars would quarrel with my use of the word “prophetic” here. Lewis distinguishes among prophetic dreams, precognitive dreams, and traditional divinatory dreams. Lewis, supra note 11, at 192. Precognitive dreams usually involve some trivial future event over which the dreamer has no control. Id. In contrast, a prophetic dream also tells the future, but the events “relate to important areas of life. As with the Hebrew prophets, there is also some sense that a prophetic dream gives one a chance to actually change the future, as if the dreamed events do not have to happen or can be modified in some way.” Id. Divinatory dreams are more symbolic in content. Id. For example, if one dreams of falling off an office building, and the next week loses his job, the dream might be said to have predicted the firing. Id. Divinatory dreams require symbolic interpretation, and thus are distinguished from most other precognitive dreams. Id.

\textsuperscript{158}. See State v. Vas, 687 A.2d 1295 (Conn. 1997).

\textsuperscript{159}. Id. at 1297, 1300.

\textsuperscript{160}. Id. at 1297.

\textsuperscript{161}. Id.

\textsuperscript{162}. Id.
Wild Dreamers

upset and accused his six-year-old daughter of not informing him of the mother’s impending marriage.\(^{163}\) When Joann Daigle-Moylan came to pick up Ayla, she argued with Vas, and he ended up pointing a gun at all three of them—the mother Joann, Ayla, and Lannan.\(^{164}\) Then, according to the appellate court’s recitation of the facts in \textit{State v. Vas},\(^{165}\) Joann barricaded herself in an office. Lannan initially approached [Diego Vas] to take the gun away, but then changed her mind and turned away from him. As she turned, [Vas] fired several shots and ran from the building . . . . Daigle-Moylan came out of the office to find that Lannan and Ayla had been shot. Lannan was severely wounded but eventually recovered from her wounds. Ayla died as a result of wounds to the head. [Diego Vas] was arrested in New York City the next day.\(^{166}\)

He was charged with one count of murder and two counts of attempted murder, and was ultimately convicted.

At trial, the defense put on a witness, Thomas Juko, who testified of the defendant’s loving relationship with Ayla, describing how he had regularly observed the defendant and Ayla interacting at an eating establishment that he frequented. The state’s attorney then asked Juko, “[W]ere you aware that the child—near the time of her death—was dreaming that her father was killing her?”\(^{167}\)

Defense counsel for Diego Vas objected, and the judge instructed the jury to disregard the question.\(^{168}\) Defense counsel then moved for a mistrial; the motion was denied, and again in its charge to the jury, the judge told the jurors to disregard the prosecutor’s question about the victim’s dream that her father was killing her—just in case they had forgotten about it altogether.\(^{169}\) On appeal, most of the opinion in \textit{State v. Vas} was about the trial court’s instruction to the jury about the defense of mental disease or defect;\(^{170}\) regarding the jury’s hearing about Ayla’s dreams of murder, the appellate court concluded that the defendant had

\(^{163}\) \textit{Id.}.
\(^{164}\) \textit{Id.}
\(^{165}\) 687 A.2d 1295 (Conn. 1997).
\(^{166}\) \textit{Id.} at 1297.
\(^{167}\) \textit{Id.} at 1300.
\(^{168}\) \textit{Id.}
\(^{169}\) \textit{Id.}
\(^{170}\) \textit{Id.} at 1297–1300.
“failed to show that the state’s question to Juko was so prejudicial that it deprived him of a fair trial.”\textsuperscript{171}

Unlike O.J. Simpson, who was a famous dreaming man, this wild dreamer was just a little girl who lived for a short time somewhere in central Connecticut—unknown to the world except for her brief posthumous appearance in the \textit{Atlantic Reporter}. Having no more information about her, the storyteller in me feels driven to invent some context for her murder. Over and over on my mind’s movie screen, I direct various scenes of this sad tale, not always with satisfaction—the only script I have are the few sentences written in a lower court’s opinion.

Here is the first take. \textit{Scene One}: Ayla and her father, kidding around in a diner, Diego making funny faces at his little girl behind a plastic menu. Ayla, shy, but smiling, pleased at the attention. Thomas, sitting in the booth next to the window, drinking his coffee, black, two sugars, watching the two make fun, envious of the love and affection that flowed between father and daughter. \textit{Scene Two}: Diego sitting on a gray naugahyde couch in the waiting room of the Child Protection Council offices in Danielson, leafing nervously through a \textit{Parents Magazine} as he waits for Joyce to call him into the Visiting Room. (I am plagued by weaknesses in the plot. Why is the father in the diner with Ayla if he only has supervised parenting privileges? Did the scenes described by Thomas predate some shameful conduct that forced the supervised parenting on Diego? And who is this guy, Thomas? Is this the best witness the defense could muster, a guy who has watched Diego interact with his daughter at some undisclosed time in “an eating establishment that he frequented”?) \textit{Scene Three}: Joyce—I see her as an efficient woman of about thirty, slightly overweight, sincere, concerned. She is pouring over Ayla’s file and frowns when she encounters an entry from a caseworker: “Ayla reports dreaming often that her father was killing her.”

Maybe I have the script all wrong. Did Ayla tell her dreams to her caseworker, or to her mother? Or was there a grandmother somewhere in the background, the kind of grandmother a six-year-old could confide in—who might listen to a little girl’s nightmares? Perhaps her mother, Joann, was too busy with the new love of her life to listen. One thing is for sure: Ayla did not tell Thomas Juko her dreams. Why would he have had any reason to know what Ayla was dreaming? Dream-telling in our

\textsuperscript{171}. \textit{Id.} at 1300–01.
culture is an intimate exchange, not one to take place over the top of a booth in a diner, between strangers. No one would believe my first scene, if I had directed Ayla to lean over suddenly and remark from the side of the menu to Thomas Juko, “Hey, mister, you might think this looks like a lot of laughs, this warm exchange between me and my father, but you ought to know that at night, when I am alone and asleep, I dream another truth: that the man across this formica table top is killing me.”

How did those lawyers in the courtroom feel about Ayla’s dream? We know the trial judge ruled that Ayla’s dream of being murdered by her father was inadmissible. We also know the trial judge told the jury to disregard the question, both when it was introduced, and in the charge to the jury at the end of the trial. But even for an off-hand statement in a cross-examination, struck from the record, the mention of Ayla’s dream must have raised some eyebrows in the courtroom. Why else would the defense argue on appeal that the reference to her dream warranted a mistrial? Defense counsel must have believed that something very harmful had happened to Diego in that one slip of the prosecutor’s tongue.

And how slipful was that question? Surely the prosecutor must have made a tactical decision to inflict that harm. He intentionally asked Thomas Juko whether he knew that Ayla had dreamed of her father killing her. No doubt his rationale for asking the question was to undermine Juko’s credibility, to show that Juko lacked knowledge of the defendant and his relationship with his daughter, to show that he only knew the partial truth—the superficial image of the playful father and laughing daughter in the diner. He meant to say to Juko and to the jurors: that same child—laughing and loving, so you say—that same child cried out at night, convinced in her dreams that her father was going to kill her. But there are rules about such cross-examination, and the prosecutor must have known that he had broken them—that he had asked a question that went far beyond testing Juko’s knowledge of the defendant and his situation.172

172. I can only assume Juko was a character witness for the defendant. For example, with “specific act” cross-examination of a reputation witness, it is permissible to ask the reputation witness whether he knew that the defendant had committed prior bad acts, not to show that he had generally a bad character, but to test the knowledge and credibility of the witness. See GLEN WEISENBERGER & JAMES J. DUAHÉ, FEDERAL RULES OF EVIDENCE: RULES, LEGISLATIVE HISTORY, COMMENTARY AND AUTHORITY § 405.4, at 129 (2001). The government has to show that there was a good-faith factual basis for the incidents raised on cross-examination, and that they are relevant to the character traits at issue. Id. Plus, the prior bad acts must have been likely to have
And where would a six-year-old girl come up with a dream that her father was killing her? And when she cried out at night, was there someone who came into her room to lead her back to a safer reality? Did she have a safer reality? It seems not. Who heard about those dreams? Why was there no state of alarm declared when a six-year-old child tells someone that she dreams of her father killing her? Why was Diego Vas allowed to enter the child protection office with a gun? Why did her mother barricade only herself in an office when Vas pulled out his gun, and not her daughter as well? Who was protecting that child? How could they not have suspected she was in danger? To my mind, a six-year-old child who dreams that her father is murdering her might well be in danger.\textsuperscript{173} It does not strike me as a universal dream that a healthy, secure six-year-old might have. At a very minimum, the dream reveals that Ayla felt she might be in danger, and that ought to have been enough to put them on alert.

I wish someone in her family or the social work system had listened to Ayla’s dreams. She needed and deserved better protection. But I would not have wanted the jury trying Ayla’s father for her murder to have heard about her dreams. As dreams of a living six-year-old girl narrated to her family or social worker, they are indications of her state of mind and cries for help. But as dreams of a dead six-year-old girl mentioned in a prosecution for her murder, we are again plagued by problems of hearsay—forgetting for a moment the issue of relevance. Just as with O.J. Simpson’s dreams of killing his wife, there is the concern about a hidden layer of hearsay in this case: that Ayla’s dream was a prophecy, and originated from another declarant, one unseen, possibly one divine. No one is going to argue that point, but I assure you, some members of the jury might be thinking about it.

\textsuperscript{173} Unfortunately, Ayla is not alone in dreaming of her victimization. In a study of 300 dreams of children ages two through twelve years, there was a high level of physical aggression, and usually the dreamer was the victim. Patricia Garfield, Your Child’s Dreams 46–47 (1984). The younger the child, the more frequently they reported becoming the victim of someone’s aggression in a dream, indicating that children are “more anxious about victimization than adults are.” \textit{Id.} “As adults, females are more often victims of violence in their dreams than men are, and it is most often the dream men who commit these crimes; it is a male-dominated dream world.” \textit{Id.} at 47.
Wild Dreamers

Every time I tell Ayla’s sad story to someone, when I get to the part about the prosecutor’s cross-examination, where he asked Juko, “Were you aware that the child—near the time of her death—was dreaming that her father was killing her?”—the listener makes a low moan. The same low moan that I made on the first reading of State v. Vas. My guess is that some members of that jury also let out a low moan when they heard the question—and that the prosecutor was counting on that. A moan is defined as an “audible expression of sorrow or suffering.” Moans do not emanate from rational thought, from inferences of knowledge carefully drawn from sense data. Moans emanate from the primitive self, perhaps from the medieval self, from the self that recognizes knowledge of a different sort—knowledge that comes from the bones, from the heart, from the premonitions of a dead child who cried out to no one in the night.

And I am confident of this: the prosecutor knew that he had no theory of admissibility for Ayla’s dreams. Again, I can hear him thinking, even if I could get the statement in as an exception to the hearsay rule, there is the more foundational issue of relevance. Ayla’s dream is even more problematic than O.J.’s because at least his dreams could be deemed relevant on a rough and ready version of a Freudian wish-fulfillment theory. The state of mind of the defendant relates to the issue of mens rea. Here, when the dreamer was the victim and not the perpetrator, there is no “state of mind” theory to justify the dream’s introduction into evidence. Ayla is not on trial for the murder of her father. Hence, there is no theory of self-defense making the victim’s state of mind relevant—no argument before the jury that Ayla was afraid of her father, hence justifying an act of violence against him. All the prosecutor had in his arsenal was a lame effort to impeach the witness, Thomas Juko, on how well he knew the defendant, on how competent he was to state an opinion about the defendant’s warm relationship with his daughter.

Under a rationalist scheme, there is simply no theory to make Ayla’s dreams relevant. If the prosecutor wanted to stay faithful to the empirical episteme, how could he argue that Ayla’s dreams make it more likely than not that the defendant killed the victim? But while not legally relevant, her story is compelling—and the prosecutor knew that too: Ayla dreamed of being killed by the defendant; Ayla was killed by the defendant. Her murder gave the dream context and meaning; it suddenly became prophetic, the first chapter of a story. The structure of the

174. WEBSTER’S NEW UNIVERSAL UNABRIDGED DICTIONARY 1153 (2d ed. 1986).
narrative began to hold sway in the mind of the listener. The dream foreshadowed the last chapter of Ayla’s story; it set into stone the chain of causal relations that must inexorably lead to Ayla’s tragic end. Hearing about that dream would be highly prejudicial to the defendant in his trial for Ayla’s murder. For many dream interpreters, he would not stand a chance.

It is time for a Jungian interlude.

D. A Jungian Interlude—Ayla’s Dream and the Archetype of the Murdering Father

Here is an understatement: Carl Jung was not an empiricist.\textsuperscript{175} He might have embraced the epistemological stance for the conscious mind—that we may only claim to “know” things experienced with one’s sensory apparatus. But for Jung, to focus solely on the functioning of the conscious mind reflects too narrow a view of the human psyche.\textsuperscript{176} He believed that the unconscious was capable of attaining a more profound awareness of matters beyond our perception, and that at least part of the psyche is “not subject to the laws of space and time.”\textsuperscript{177} “True,” he

\textsuperscript{175} Freud was Jung’s elder by nineteen years, and at one time early in their relationship, he confided to Jung that he was adopting him “as an eldest son, anointing him as successor and crown prince.” Joseph Campbell, \textit{Introduction} to \textit{The Portable Jung}, at vii, xvii (Joseph Campbell ed., 1976) (internal quotations omitted). Freud and Jung eventually had a professional and personal falling out that centered around Jung’s rejection of some of Freud’s ideas, as well as Jung’s openness to alternative epistemes, manifested in his interest in precognition and parapsychology, which Freud dubbed: “Sheer nonsense!” Jung characterized Freud’s response to his views in terms “of so shallow a positivism that I had difficulty in checking the sharp retort on the tip of my tongue.” \textit{Id.}

\textsuperscript{176} According to Jung, psychic life does not consist “only of self-evident matters”; if it did, then “we could content ourselves with a sturdy empiricism.” \textit{Carl Jung, The Stages of Life}, in \textit{The Portable Jung}, supra note 175, at 3, 4. Rather, our psychic processes are “made up to a large extent of reflections, doubts, experiments, all of which are almost completely foreign to the unconscious, instinctive mind of primitive man.” \textit{Id.} As we move out of childhood, we abandon nature “and are driven to consciousness. There is no other way open to us; we are forced to resort to conscious decisions and solutions where formerly we trusted ourselves to natural happenings.” \textit{Id.} But our conscious thought is only the tip of the iceberg. We also unconsciously think “in primordial images, in symbols which are older than historical man, which are inborn in him from the earliest times, and, eternally living, outlasting all generations.” \textit{Id.} at 21. These images still make up the groundwork of the human psyche. It is only possible to live the fullest life when we are in harmony with these symbols; wisdom is a return to them. It is a question neither of belief nor of knowledge, but of the agreement of our thinking with the primordial images of the unconscious.

\textit{Id.}

\textsuperscript{177} \textit{Carl Jung, Memories, Dreams, Reflections} 304 (Richard & Clara Winstons trans., 1963) [hereinafter \textit{Memories}].
argued, “the unconscious knows more than the consciousness does; but it is knowledge of a special sort, knowledge in eternity, usually without reference to the here and now, not couched in language of the intellect.”178 Jung believed that we get a glimmer of these “irrepresentable realities” when we make mathematical statements about infinity, or when we engage in dream analysis, or when we have experiences that evidence proof of life after death.179

Jung’s autobiography is full of stories with premonitions, precognitive dreams, intuitions, and visions.180 In the Prologue, Jung warns the reader that his memoirs are shy on external events and outward circumstances.181 Rather, he could only understand himself “in the light of inner happenings.”182 Recollection of the outward events in his life had “largely faded or disappeared. But my encounters with the ‘other’ reality, my bouts with the unconscious, are indelibly engraved upon my memory. In that realm there has always been wealth in abundance, and everything else lost importance by comparison.”183

Jung believed that the unconscious had two discernible layers.184 The first layer he dubbed the personal unconscious; it was specific to the individual, consisting of forgotten bits of the past, or repressed thought from which the consciousness had withdrawn, or of sense impressions that “never had sufficient intensity to reach consciousness but have

178. Id. at 311.
179. Id. at 304–11.
180. In his autobiography, Jung described numerous paranormal experiences. Here is an example: Several nights after the sudden death of a friend, Jung awoke from sleep, and felt that the friend was in his room. Id. at 312. The friend was standing at the foot of Jung’s bed, and beckoned him to go with him. Id. Jung asked himself: could this be a fantasy or is my friend really there? Id. Then, in a cool, rational way, Jung reasoned: “Proof is neither here nor there! Instead of explaining him away as a fantasy, I might just as well give him the benefit of the doubt and for experiment’s sake credit him with reality.” Id. (internal quotations omitted). After that thought, his friend “went to the door and beckoned me to follow him. So I was going to have to play along with him! That was something I hadn’t bargained for. I had to repeat my argument to myself once more. Only then did I follow him in my imagination.” Id. Jung followed the dead friend who led him to his former library, climbed a stool, and pulled down the second of five books with red bindings that stood on the second shelf from the top. Id. Then the vision was over. Id. When Jung went to visit his friend’s widow the next day, he asked if he could look something up in the dead man’s library, and there was the stool, and on the second shelf, there were the five books with red bindings. Id. at 312–13. The books turned out to be the novels of Emile Zola. Id. at 313. Jung climbed onto the stool, and read the title of the second volume: The Legacy of the Dead. Id.
181. Id. at 5.
182. Id.
183. Id.
184. CARL JUNG, The Structure of the Psyche, in THE PORTABLE JUNG, supra note 175, at 23, 38.
somehow entered the psyche.” The second layer, called the collective unconscious, is not individual but common to all men; it is the “ancestral heritage of possibilities of representation.” The collective unconscious “so far as we can say anything about it at all—appears to consist of mythological motifs or primordial images, for which reason the myths of all nations are its real exponents.”

There are myth motifs in all cultures that Jung calls “archetypes.” For example, many myths and fables portray rivers as containing spirits, and thunder as the voice of an angry god, with lightning as “his avenging missile.” There is a “universal hero myth” that depicts a “powerful man or god-man who vanquishes evil in the form of dragons, serpents, monsters, demons, and so on, and who liberates his people from destruction and death.” Even human families consist of archetypes: the father, the earth mother, the innocent child—all these myth-motifs are eternally and universally repeated in all cultures. In Jung’s words:

185. Id.
186. Id. Mythology and primordial images are the exponents of the collective unconscious. The collective unconscious is the totality of all archetypes, is the deposit of all human experience right back to its remotest beginnings. Not, indeed, a dead deposit, a sort of abandoned rubbish-heap, but a living system of reactions and aptitudes that determine the individual’s life in invisible ways—all the more effective because invisible. Id. at 44.
187. Id. at 39. Under Jung’s theory, there are three psychic levels: (1) the consciousness; (2) the personal unconscious (which belongs individually to the person and consists of contents that were forgotten because they had lost their intensity, or were repressed, or were never sufficiently intense to have reached consciousness); and (3) the collective unconscious. Id. at 38. The collective unconscious does not just belong to the individual, but is “common to all men, and perhaps even to all animals, and is the true basis of the individual psyche.” Id.
188. CARL G. JUNG ET AL., MAN AND HIS SYMBOLS 68 (1964) [hereinafter MAN AND HIS SYMBOLS].
189. Id. at 85.
190. Id. at 68.
191. According to Jung, the archetypes that have the “most frequent and the most disturbing influence on the ego” are the shadow, the anima, and the animus. CARL JUNG, PHENOMENOLOGY OF THE SELF, in THE PORTABLE JUNG, supra note 175, at 139, 145. The anima refers to personality traits regarded as feminine; they are often repressed into the unconscious in a male psyche. Id. at 151. The parallel structure in the female psyche is the animus. Id. The ego represents the individual’s sense of personal self; the shadow is a kind of unconscious “counterego.” Id. at 145. It is the negative, socially undesirable part of the self, and becoming conscious of one’s shadow cannot be achieved “without considerable moral effort.” Id. It involves “recognizing the dark aspects of the personality as present and real. This act is the essential condition for any kind of self-knowledge, and it therefore, as a rule, meets with considerable resistance.” Id.
Wild Dreamers

The deposit of mankind’s whole ancestral experience—so rich in emotional imagery—of father, mother, child, husband and wife, of the magic personality, of dangers to body and soul, has exalted this group of archetypes into the supreme regulating principles of religious and even of political life, in unconscious recognition of their tremendous psychic power.  

These archetypes, together with instincts, form the collective unconsciousness.  

Instincts are unconscious processes that are inherited, and occur uniformly and regularly among all people.  

For example, the reproduction process is instinctual.  

It is not learned behavior, but automatic.  

Intuition is sometimes referred to as an “instinctive” act of comprehension, but Jung points out that intuition is a process analogous to instinct, with the difference being that “instinct is a purposive impulse to carry out some highly complicated action, intuition is the unconscious, purposive apprehension of a highly complicated situation.”  

Intuition is one of the basic functions of the psyche, and Jung calls it “perception of the possibilities inherent in a situation.”  

Jung agreed with Freud that dreaming was a direct expression of the unconscious, but he rejected Freud’s insistence that the content of the repression was invariably a sexual trauma.  

Indeed, Jung’s entire

---

192. JUNG, supra note 184, at 43.
193. Id. at 44.
194. CARL JUNG, Instinct and the Unconscious, in THE PORTABLE JUNG, supra note 175, at 47, 50.
195. Id. at 50–51.
196. Id. at 50.
197. Id. at 51.
198. JUNG, supra note 184, at 26 (original emphasis omitted). Jung did ascribe to an empirical epistemology for conscious knowledge. “Consciousness seems to stream into us from outside in the form of sense-perceptions. We see, hear, taste, and smell the world, and so are conscious of the world.” Id. at 25 (original emphasis omitted). From sensations, we think about what things might be, and we have feelings towards those things. Id. at 25–26. But apart from sense-perception, thinking, and feeling, there is intuition. Id. at 25. About intuition, “language shows a regrettable lack of discrimination.” Id. at 26.
199. MEMORIES, supra note 177, at 147. Jung could not accept Freud’s identification of the content of repression as invariably sexual trauma. In 1910, Freud and Jung had a famous encounter at the Second Congress of the Association of Psycho-Analysis where Freud insisted, even against organized opposition, that Jung should be appointed the permanent president. He urged Jung to “promise me never to abandon the sexual theory. That is the most essential thing of all. You see, we must make a dogma of it, an unshakable bulwark.” Campbell, supra note 175, at xviii. When Jung asked him, a bulwark against what, Freud replied, “‘Against the black tide of mud’—and here he hesitated for a moment, then added, ‘of occultism.’” Id. at xix. Jung was alarmed at the use of the
approach to dreams was pragmatic and open-minded; he aggressively sought to ensure that no preconceived theory about the meaning of dreams intruded upon their interpretation. He would protest: “I have no theory about dreams. I do not know how dreams arise. And I am not at all sure that my way of handling dreams even deserves the name of a ‘method.’” Dreams were just one of many psychic phenomena, and Jung exhorted others to “bear in mind that there is no simple and generally known theory of psychic phenomena, neither with regard to their nature, nor to their cause, nor to their purpose.” Theories like Freud’s wish-fulfillment and sexual repression were valuable points of view, but they did not do “anything like justice to the profundity and richness of the human psyche.”

Consistent with Jung’s rejection of dogmatism, he took the position that “it should be an absolute rule that every dream, and every part of a dream, is unknown at the outset, and to attempt an interpretation only after carefully taking up the context.” Each dream is embedded in a web of associations that belong and are known to only the individual dreamer, yet the dreamer might also include images from our shared collective experience in the form of archetype. Jung found that sometimes the dreams of his patients contained images with which they

word dogma, an “undisputable confession of faith” contrary to scientific judgment. Id. But he was also upset by Freud’s fear of “occultism.” Id. Jung wrote:

“I knew that I would never be able to accept such an attitude. What Freud seemed to mean by ‘occultism’ was virtually everything that philosophy and religion, including the rising contemporary science of parapsychology, had learned about the psyche. To me the sexual theory was just as occult . . . a hypothesis that might be adequate for the moment but was not to be preserved as an article of faith for all time.”

Id. at xix (quoting Jung).

200. Jung disagreed with Freud that the dream is a “‘facade’ behind which its meaning lies hidden—a meaning already known but maliciously, so to speak, withheld from consciousness.” MEMORIES, supra note 177, at 161. To Jung, dreams are a part of nature, which harbors no intention to deceive, but expresses something as best it can, just as a plant grows or an animal seeks its food as best it can. . . . Long before I met Freud I regarded the unconscious, and dreams, which are its direct exponents, as natural processes to which no arbitrariness can be attributed.

Id. at 161–62.


202. Id.

203. Id. at 299.

204. Id.

205. Id. at 310.

206. Id. at 315.
were unfamiliar, but which reflected symbols that could be found in the world’s mythological systems. The analyst’s job was to take these two layers of meaning—the personal and the collective or universal—and relate them to the dream’s situation.

For Jung, the dream served two functions. First, it compensated for internal imbalances in the psyche, and second, it assisted in what Jung called the “individuation process,” which Jung translated as “coming to selfhood” or “self-realization.” The dream’s purpose was to communicate something to the conscious mind. Thus, the dream had a pragmatic function and was future-oriented. The unconscious had its own intelligence, and the dream was a tool of revelation—something for the conscious mind to learn from and use in its ongoing quest for self-realization and individuation.

How would Carl Jung interpret Ayla’s dream of her father murdering her? At one time in his life, Jung would not have deemed her precognitive dream a prophecy. In his earlier writings, he had this to say about dreams that “tell the future”:

The occurrence of prospective dreams cannot be denied. It would be wrong to call them prophetic, because at bottom they are no more prophetic than a medical diagnosis or a weather

207. See, e.g., id. (relating account of a young man’s dream of performing fellatio on himself and noting that it was an exact parallel to an ancient Hindu myth of creation).
208. Id. at 306–07 (“The task of the analyst is to differentiate these two layers and to relate them relevantly to the conscious life situation of the dreamer. Such a skill requires of the analyst that he or she be well read and deeply familiar with the images and motifs of religion, mythology, folklore, fairy tale, drama, and the arts. When such images and motifs appear in dreams, the archetypical layer is impinging upon the personal ‘story’ of the dreamer.”).
209. Id. at 302. Jung also theorized that the dream could be regarded as compensation or correction for unbalanced conscious thoughts. The nature of psychic process is a “dialectical interplay between pairs of opposites, most especially that between ego consciousness and the unconscious. Consequently, the dream is understood as the primary mode by which the unconscious expresses in symbolic form a balancing or homeostatic reaction to the one-sided position of the conscious attitude.” Id. Jung interpreted a dream that he had that presaged his forthcoming break with Freud. He had dreamed that Freud was a “peevious official of the Imperial Austrian monarchy, as a defunct and still walking ghost of a customs inspector.” MEMORIES, supra note 177, at 164. He postulated, “But it was possible that the dream could be regarded as a corrective, as a compensation or antidote for my conscious high opinion and admiration. Therefore the dream recommended a rather more critical attitude toward Freud.” Id.
210. MEMORIES, supra note 177, at 383.
211. Greene, supra note 201, at 304.
212. Id. at 304–05.
213. Id. at 305.
214. See supra Part II.C.
forecast. They are merely an anticipatory combination of probabilities which may coincide with the actual behavior of things but need not necessarily agree in every detail.215

He might have said: Ayla’s dream was like a bit of dreaming intuition. She was looking at the possibilities in the situation—and had concluded her father just might murder her.

But later in his life, Jung seemed more open to the idea of a dream as prophecy. He admitted there were documented cases of what he referred to as “spontaneous foreknowledge,” sometimes occurring in precognitive dreams.216 Because a rationalist only believes in the reality of the phenomenal world, Jung explained, he must reject the existence of such para-psychological experiences.217 But if those phenomena occur, and Jung later believed they did, then the “rationalistic picture of the universe is invalid, because incomplete.” A complete picture of the world would require the addition of still another dimension; only then could the totality of phenomena be given a unified explanation.218 So maybe if we asked Jung later in his life, he might be willing to say that Ayla had been prophetic—and that time might go in both directions.220

Either way, Jung would surely be willing to say that Ayla’s dream was informed by the collective unconscious, the ancient, primitive store of symbolic imagery and mythology that transcends particularity. Myths contain several parental archetypes. The earth mother, for example, is one, and her devotion to her child is laudable and comforting.221 But the

215. Greene, supra note 201, at 305 (quoting Jung).
216. MEMORIES, supra note 177, at 304; see also supra note 157.
217. MEMORIES, supra note 177, at 305.
218. Id.
219. Id.
220. This is a subject that has generated a great deal of scientific literature. See, e.g., P.C.W. DAVIES, THE PHYSICS OF TIME ASYMMETRY (1977); PHYSICAL ORIGINS OF TIME ASYMMETRY (J.J. Halliwell et al. eds., 1994); H. PRICE, TIME’S ARROW AND ARCHIMEDES’ POINT (1996).
221. Demeter’s story is told in the earliest of the Homeric poems. Demeter, the goddess of the

- corn, the harvest, the summer time, only had one daughter, Persephone, who was the maiden of spring. EDITH HAMILTON, MYTHOLOGY 49 (1940). Lured by the bloom of the narcissus, Persephone strayed too far and was taken by the lord of the dark underworld where she dwelled among the shadowy dead. Id. at 50. Demeter grieved her loss so much that nothing would grow on earth, and there was famine in the land. Id. at 52. Zeus intervened, and sent Hermes to the underworld to fetch Persephone. Id. The lord of the underworld made Persephone swallow a pomegranate seed, “knowing in his heart that if she did she must return to him.” Id. Persephone returned to her mother, and the fields were once more abundant with fruit and greenery, but every year, for four months, Persephone had to return to the underworld. Id. at 53–54.

After the lord of the dark world below carried her away she was never again the gay young creature who had played in the flowery meadow without a thought of care or trouble. She did
Wild Dreamers

myths manifesting the collective unconscious sometimes reveal the shadow, the darker side of the human psyche—including the parent’s potential for selfishness and jealousy. Perhaps it was the archetype of Cronos that Ayla’s unconscious was drawing upon—the Titan father who destroyed his young by swallowing them whole when he learned that one of them would someday steal his throne. Sometimes the parent is willing to sacrifice his child if it benefits him in some way, or if he feels threatened, or if someone else wants to take the child from him. Anyone who has ever spent any time in family court knows this to be true—that children are often used as pawns by their parents in bitter battles between husband and wife. Entitlement to the child trumps concern for the child. Children can even get murdered that way—if I cannot have her, then neither can you. Ayla sensed that the father who gave her life believed that he had the power to destroy her if things did not go his way. It is an old story, Jung would say, and no quirk of fate that Ayla would have access to it in her sleep—even before it happened. About this I am certain: Jung would have exhorted Ayla’s protectors to be attentive to her dreams.

indeed rise from the dead every spring, but she brought with her the memory of where she had come from; with all her bright beauty there was something strange and awesome about her. She was often said to be,”the maiden whose name may not be spoken.”

Id. at 54.

222. Cronos, or Saturn in Latin, was one of the Titans, often called the Elder Gods. Id. at 65. The Titans were for many ages supreme in the universe, and of enormous size and strength. Id. The most important Titan was Cronos who ruled over the other Titans until his son, Zeus, vanquished him and seized the throne for himself. Id. at 66–67. Cronos had himself wounded his own father, and he was more than a little sensitive about being dethroned by his own son. Id. at 65. Cronos had learned that one of his children would one day dethrone him, so “he thought to go against fate by swallowing them as soon as they were born.” Id. After his sister-queen, Rhea, gave birth to Zeus, she had him secretly removed to Crete, and gave her husband a stone wrapped in swaddling clothes that he promptly swallowed. Id. at 65–66. When Zeus grew up, he forced Cronos to disgorge the stone, along with his five siblings. Id. at 66. Eventually, after a terrible war, Zeus and his siblings vanquished Cronos and his brother Titans. Id. at 66.

223. These stories show up in the newspaper from time to time. Here is an example: Woman in Custody Case Kills Herself and Her Child, Police Say was the caption for a brief story in the Metro Section of the New York Times. See Bruce Lambert, Woman in Custody Case Kills Herself and Her Child, Police Say, N.Y. TIMES, July 25, 2003, at B5. The bodies of a thirty-four-year-old woman and her two-year-old daughter were found in a Long Island home the day before the woman was due to appear in Nassau County Family Court in a child custody suit; the deaths appeared to be a murder-suicide, based on the note left behind. Id. The girl was found face down in the bathtub, underwater. Id. The last sentence of the newspaper article read, “A complaint about the family’s situation was reported to the county’s child protection agency earlier this year, but an investigation determined that the report was unfounded, the county social services commissioner, Bob Sherman, said last night.” Id.

224. See id.
But ours is not a culture where dreams matter much. It is time to end Ayla’s story with a discussion of the Iroquois—a people who had a dreaming culture. Sad but true, we have to move back in time, before the Europeans set foot on North America, when the belief system of the Iroquois was still intact. Ayla might still be alive today if she had belonged to a dreaming culture like that of the Iroquois.

E. Concluding Ayla’s Story—The Dreaming Culture of the Iroquois

Much of what we know about the Iroquois comes from first-hand accounts of Europeans, many of them Jesuit priests who settled in the northeast during the seventeenth century. From their letters and diaries, we have learned that the Iroquois took dreaming very seriously. Dream-sharing was a common practice among the Iroquois. Sometimes if the dreamer could not remember or fully understand a dream, a shaman possessing visionary powers would be

---

225. Even in the late twentieth century, the dreaming cultures of North American native people were much stronger than non-native populations:

Not only did a group of Cree Indians from Alberta recall significantly more dreams and feel more alert on awakening than did non-Indian control subjects, but they made more attempts to understand and analyze their dreams. They were also significantly more likely to contact the people who appeared in their dreams—thus maintaining the traditional performative utilization of dreams.

HUNT, supra note 1, at 85.

Lydia Degarrod provides a fascinating account of how the Mapuche in southern Chile use dream narratives to aid in the formation of a distinct ethnic identity, and to help establish a cultural resistance to oppression and domination. See generally Lydia Nakashima Degarrod, Mapuche: Dream Interpretation and Ethnic Consciousness, in NEW DIRECTIONS IN ANTHROPOLOGY (Prentice Hall CD-ROM, Apr. 14, 2003).

226. “Iroquois” is a term used to include the Five Nations of the Iroquois and the Huron. Donald Patrick St. John, The Dream-Vision Experience of the Iroquois: Its Religious Meaning 6 (1981) (unpublished Ph.D. dissertation, Fordham University) (on file with author). The Iroquois lived in the northeast prior to the beginning of the Christian era, probably initially in small, mobile bands, subsisting on hunting, fishing, and gathering. Id. at 6–7. They became settled by 500 C.E., and after 1000 C.E., the Iroquois moved from villages along rivers and banks to more permanent settlements on hill tops, consolidating in larger tribal units. Id. at 7–8. The League of the Iroquois was formed in the middle of the fifteenth century. Id. at 8.

227. Id. at 24. Most of the diaries and letters were written by French Jesuit missionaries who lived among the native people of the Northeast in the seventeenth and early eighteenth centuries. Id. Considered the “single largest source of data on seventeenth century Iroquoian life,” the “Jesuit Relations,” consisting of the written reports, journals, and correspondence left by the missionaries, resulted in “an admirable account of Indian life and customs, even considering the conditions of the time and the fundamental differences between the Indians and the Europeans. It is mainly to the Relations that we owe our knowledge of the role of the dream-vision experience in Iroquoian society.” Id.

228. Id. at 62.
Wild Dreamers
called in to divine its meaning. During the great annual Midwinter Festival, and at other times of the year if a healing ceremony was called for, the Iroquois engaged in the onoharoia, or the dream-guessing ceremony. At this ceremony, old and new dreams would be told, and the community would respond by performing various rituals designed to “disarm the spirits who might bring disease to the people across the ice-bridge from the old year.” There were expert “interpreters,” or artemidores, who “tried to make sense out of the confused images in order to create a coherent dream sequence.” At times, “communal dramatization of the dream might be demanded” by the dreamer. The fear was that “great peril would come to the nation” if the dream was not acted out.

The Iroquois believed that “sympathetic communication can exist between two psychophysical beings in the sacred cosmos”; sometimes that communication took place in dreams. One of the closest relationships in this sacred cosmos was between a human being and his guardian spirit, often an animal spirit guide. With respect to animal spirit guides, it was crucial for the Iroquois, and in particular the Huron who depended less on agriculture, to be successful in their winter hunts. Success depended on “the maintenance of a proper relationship with those sacred beings who watched over the animals.” Animals were created by the Master of Life, who instructed them that they must give themselves over to humans so that the latter might live; animals therefore needed to be “related to both their spiritual and material dimensions.” Through the dream, humans encountered either the “tutelary spirit of an animal species or the spirit (‘soul’) of a particular animal.” Such an encounter was seen as a fortuitous sign, and usually

229. Id.
230. Id. at 64–65.
231. Id. at 66.
232. Id. at 68.
233. Id.
234. Id.
235. Id. at 51.
236. Id. at 54–55.
237. Id. at 49.
238. Id.
239. Id.
240. Id.
guaranteed success in hunting. 241 Dream visitors, such as the animal spirit guides, came from a dimension beyond the physical plane. 242

For the Iroquois, dreams sometimes foretold the future. 243 At other times, the dream relayed vital information to the dreamer—often about his health or survival. 244 Dreams were essential to healing. 245 The Iroquois understood the intimate relationship between the spirit and the body—how the former, if not protected, could poison the latter. As one Jesuit priest reported, the soul makes its wishes

"known by means of dreams, which are its language. Accordingly, when these desires are accomplished, it is satisfied; but, on the contrary, if it be not granted what it desires, it becomes angry, and not only does not give its body the good and happiness that it wished to procure for it, but often it also revolts against the body, causing various diseases, even death." 246

Another seventeenth century Jesuit observer wrote that for the Iroquois, disease comes from the "mind of the patient himself, which desires something, and will vex the body of the sick man until it possesses the thing required." 247 He added:

"[T]hey think that there are in every man certain inborn desires, often unknown to themselves, upon which the happiness of individuals depends. For the purpose of ascertaining desires and innate appetites of this character, they summon soothsayers, who

241. Id.
242. Id. at 49–50. The Iroquois assumed a network of rights and duties among all beings: Whether one speaks of the great cosmic forces or the soul of a fish, nothing is autonomous and independent but has a place and duties "assigned" to it by the Master of Life. There is also the assumption that sympathetic communication can exist between two psychophysical beings in the sacred cosmos. In times of disharmony within the cosmic order, such communication is absolutely necessary for the restoration of balance. Sometimes a non-human person may take the initiative by appearing in a dream. Sometimes a human being through ritual will try to communicate with or influence one of the other beings in the cosmos. Id. at 50–51.
243. Id. at 69.
244. Id. at 61.
245. Id. at 58.
246. Id. at 57–58 (quoting Relation of Father Raguenau).
Wild Dreamers

as they think, have a divinely-imparted power to look into the
inmost recesses of the mind."

Moreover, the Iroquois believed that fulfillment of the frustrated
desire, either directly or symbolically, would provide relief of the
dreamer’s distress. This meant that for the Iroquois, dreams, and their
fulfillment, were a matter of public discussion and concern. About the
Seneca, one Jesuit priest wrote that these “‘people think only of
dreams, they talk about nothing else, and all their cabins are filled with
their dreams.'” Needless to say, the black-robed Jesuit priests were
worried that some Seneca might have dreams of killing them. Father
Fremin wrote:

“What peril we are in every day . . . among people who will
murder us in cold blood if they have dreamed of doing so; and
how slight needs to be an offense that a Barbarian has received
from someone, to enable his heated imagination to represent to
him in a dream that he takes revenge on the offender.”

At times, the members of this dreaming culture would get involved in
the interpretation of certain dreams—those that were thought to contain
a warning of some kind: impending death, disease, or other threats to the
people or to the peace of the community. A prophecy in a dream
might require certain rituals to ensure that the future event would not
take place. Sometimes if the dream foretold personal disaster for the
dreamer, the community could prevent the event by symbolically acting

248. Id. (quoting Relation of Father Jouvency, 1610–1613). According to Iroquois theory, there
were actually three sources of disease or bodily infirmity: from the mind of the patient himself;
from natural injuries, “such as wounds of war or physical accident; from witchcraft, by which
certain foreign articles such as balls of hair, splinters of bone, clots of blood, or bear’s teeth were
projected magically into a victim’s body.” Id.

249. There are many accounts from Jesuit priests of these public dream fulfillments:
He who has dreamed during the night that he was bathing, runs immediately, as soon as he
rises, all naked, to several cabins, in each of which he has a kettleful of water thrown over his
body, however cold the weather may be. Another who has dreamed that he was taken prisoner
and burned alive, has found himself bound and burned like a captive on the next day, being
persuaded that by thus satisfying his dream, this fidelity will avert from him the pain and
infamy of captivity and death.

Id. at 235 (quoting Relation of Father Fremin).

250. Id. (quoting Relation of Father Fremin).

251. Id.

252. Id. (quoting Relation of Father Fremin).

253. See generally St. John, supra note 226, at 54–70.

254. See id. at 69–70.
out the dream. 255 For example, one Iroquois man dreamed that he was going to be captured and tortured by the enemy. 256

To prevent this, his friends bound him and dragged him through the streets to a scaffold. Here they lit a fire in anticipation of the torturing. After he has sung his death song he was satisfied that enough of the dream has been fulfilled to prevent the actual event from occurring. 257

One early Jesuit missionary concluded that the Iroquois “speak of their dreams as of a God, they mean nothing else than it is by this means that they gain knowledge of the will of God . . . and that the doing of what they have seen in dreams is a means which contributes to the establishment of their health and of their good fortune.” 258

In many ways, the Iroquois were centuries ahead of Freud. While they had no theory of the unconscious, they recognized that dreams had a symbolic importance. There was also a public forum for the interpretation of important dreams, and a process by which to interpret them, including the assistance of a trained, expert interpreter to discover the dream’s latent meaning. 259 For the Iroquois, dreams were privileged and regarded as a source of sacred knowledge.

255. Id. at 68–69.
256. Id. at 69.
257. Id. Sometimes the dream premonitions had to be fully acted out:

One man had a dream that his house burned down. He kept insisting, to a reluctant group of elders or chiefs, that he must carry out that dream (possibly because he was afraid that he might be asleep when it actually occurred). After weighing the matter, the chiefs with great ceremony carried out the act for him.

Id. at 69–70.

258. Id. at 71 (quoting Relation of Father Milet).

259. Id. at 68. In our culture, the only trained, expert interpreters of dreams are therapists. If the dream is interpreted as a threat of violence to another person, under the rationale in Tarasoff v. Regents of the University of California, 551 P.2d 334 (Cal. 1976), the therapist may have a duty to warn the intended victim if he determines that the patient “presents a serious danger of violence to another.” Id. at 340; see, e.g., Brown v. Smith, No. X03CV000053181S, 2001 WL 1004162, at *2 (Conn. Super. Ct. Aug. 7, 2001) (imposing duty to warn on the therapist when the patient revealed, among other things, that “he had dreams of violent outbursts, mostly directed against himself but also directed to others, especially when he returns to work”). Apparently the assailant in Brown killed four co-workers at the Connecticut Lottery and then killed himself. Id. at *1; see also Logan v. Smith, No. X03CV0000531795, 2001 WL 1004188 (Conn. Super. Ct. Aug. 7, 2001); Mlynarczyk v. Smith, No. X03CV0000530968, 2001 WL 1004183 (Conn. Super. Ct. Aug. 7, 2001); Rubelmann v. Smith, No. X03CV000053180S, 2001 WL 1004237 (Conn. Super. Ct. Aug. 3, 2001).
Ayla would have fared better with the Iroquois, in a culture where all their cabins were filled with their dreams. Perhaps the Iroquois would have interpreted her dream as a warning of impending violence, threatening not only to the dreamer’s life, but to the peace of the community. Perhaps a trained interpreter would hear her dream and conclude that Ayla suffered from soul sickness, brought on by living in a damaged, fractured home where parents were at war and a child was at risk. Perhaps someone would have recognized that a six-year-old girl who dreams of her father killing her—particularly a father with a history of violence—might have something of importance to say. At the very least, there would have been a public arena in a dreaming culture like the Iroquois for her to narrate her dream and subject it to expert interpretation—and perhaps to alert the authorities to provide her with some protection.

Ayla was just born among the wrong people, at the wrong time, and in the wrong place, where dreams are not regarded as a legitimate basis for knowledge. No one was responsible for listening to a child’s dream talk. Her dreams were private matters; they did not matter. Ultimately, neither did she.

260. In one case, at least the police took a dream interpretation seriously, although this seems to be the exception to the rule. State v. McCrady, No. 97CA33, 1999 WL 4859 (Ohio Ct. App. Dec. 18, 1998). A psychic directed Mary Dye, the dreamer, to the grave site of a murder victim, where she also witnessed the defendant driving his car away from the site. Id. at *3. Mary Dye directed police to the area of the victim’s grave. Id. On appeal, the defendant argued that the trial court should have suppressed evidence gathered in the course of executing search warrants because the detective had made material omissions in the affidavits submitted to obtain those warrants, including the fact that Mary Dye had been led to the grave site on the basis of the psychic’s dream interpretation. Id. at *6. The detective found that Dye’s story was “incredible, and . . . he did not include the dream and psychic portions of her story in his affidavit in order to bolster Dye’s credibility.” Id. Moreover, Dye had been a witness to the defendant at the site, and her story was the strongest lead the detective had. Id. The court held that even if the defendant had proven the detective had been intentional or reckless in omitting allegations about the dream or the psychic, the other facts were sufficient to find the requisite probable cause for the search warrant. Id. at *7. Obviously, the psychic’s interpretation of the dream as a basis of the dreamer’s knowledge was regarded as damaging to the dreamer’s credibility.
III. THE PENULTIMATE WILD DREAMER, CARTESIAN CONFUSION, ZHUANG ZHOU’S BUTTERFLY, AND A TIBETAN BUDDHIST INTERLUDE

A. Wild Dreamer Dreams He Has Murdered a Young Woman (and He Had Not)

Glenville Smith is the penultimate wild dreamer of these meditations—the ultimate wild dreamer is another young man, Steven Linscott, who also dreamed he had murdered a young woman, and he had not. But before we turn to Smith and Linscott’s stories, let me tell you first about Descartes, a seventeenth century French philosopher, and Zhuang Zhou, a Chinese scholar and philosopher from the fourth century B.C.E.

In his First Meditation, Rene Descartes found himself sitting by the fire, staring at his hands and wondering how he could be certain of the reality of his ordinary sense perceptions when his dreams seemed to be so real. He examined his hands, and extended them, noting that he did so “deliberately and of set purpose.” His perceived reality seemed so real—he had to be awake. Then he pondered how real similar perceptions of familiar things had seemed in his dreams. Could he trust his senses? How could he know whether he was awake or in a dream? He remarked:

[...] in thinking over this I remind myself that on many occasions I have in sleep been deceived by similar illusions, and in dwelling carefully on this reflection I see so manifestly that there are no certain indications by which we may clearly distinguish wakefulness from sleep that I am lost in astonishment. And my astonishment is such that it is almost capable of persuading me that I now dream.

Descartes was one of the few philosophers in the West to consider the phenomenon of dreaming, but aside from the questions that dreams pose

261. See infra Part IV.
263. Id. at 146.
264. Id.
265. Id.
266. Id.
about the reliability of our waking senses, he had little to say about the
dreams themselves. It is from Descartes’ skepticism, however, that we
have the legacy of the scientific method. We replicate experiments over
and over again, under exactly the same conditions, and with the same
results, continually striving to reduce the unreliability of our fallible
senses.267 “All that up to the present time I have accepted as most true
and certain,” Descartes wrote, “I have learned either from the senses or
through the senses; but it is sometimes proved to me that these senses
are deceptive, and it is wiser not to trust entirely to any thing by which
we have once been deceived.”268 Descartes provided the basis for
scientific objectivity, for drawing the line between what can be claimed
as real, and what cannot.

Sometime in the fourth century before the Common Era, there was a
Chinese scholar and official named Zhuang Zhou, who meditated on the
same duality of the dreaming and waking reality as Descartes.269 In a
famous passage, Zhuang Zhou relates his dream of being a butterfly, and
of the puzzlement he felt when he awoke:

A butterfly fluttering happily around—was he revealing what he
himself meant to be? He knew nothing of Zhou. All at once
awakening, there suddenly he was—Zhou. But he didn’t know if
he was Zhou having dreamed he was a butterfly or a butterfly
dreaming he was Zhou. Between Zhou and the butterfly there
must surely be some distinction. This is known as the
transformation of things.270

267. WALTER TRUETT ANDERSON, OPEN SECRETS: A GUIDE TO TIBETAN BUDDHISM FOR
WESTERN SPIRITUAL SEEKERS 130–31 (1979). Anderson writes of Descartes’ contribution to
modern science:

Descartes, in his Discourse on Method and in Principles, worked out a set of basic concepts for
understanding the physical universe, with a radical separation of the “I,” the knower of
phenomena, from the world of observed phenomena. . . . Descartes provided the intellectual
basis for modern science: a universe of bodies moving through space according to mechanistic
laws, and a separate human intelligence capable of observing their behavior and discovering
the laws.

Id. at 110–11.

268. DESCARTES, supra note 262, at 145.

269. 1 SOURCES OF CHINESE TRADITION 103 (William Theodore de Bary & Irene Bloom eds.,
1999). Zhuang Zhou, or Zhaungzi as he is often referred to, probably lived during the late Warring
States period. Id. at 95. His work often takes the form of a conversation between real or imaginary
entities (e.g., a fictional Confucius, or a talking tree) or of parables. Id. The issues he tackles are all
highly philosophical: the nature of knowledge, the limitations of language, the relationship between
the universal and the particular. Id. at 96.

270. Id. at 103.
A man named Glenville Smith found himself caught in Zhuang Zhou’s butterfly quandary—or if you prefer to stay in the west and stare at your hands, in Descartes’ First Meditation. He had difficulties distinguishing between the real and the dreamed. These difficulties almost subjected Glenville Smith to the death penalty—he could have lost his life for being a wild dreamer. Glenville Smith had a dream about a heinous crime, and understood the dream to mean that he was the perpetrator. Smith, age twenty-eight, claimed to have psychic powers. He lived in a separate room in his girlfriend’s house in Queens; her thirteen-year-old daughter lived with them. The daughter ran away from home and was picked up by a young man from Brooklyn who suggested that they have sex. She agreed and drove back to his house where she hid out for six days. Her mother reported her missing, and the police were alerted.

A few days after the teenager was missing, Glenville Smith had a vivid and violent dream in which the girl was raped, murdered, and buried under the water in nearby Nassau County. He contacted the police, and they went out searching for her body the next day; they found nothing. Smith later took the police to six locations, and then the next day to seven other locations, but no body was ever found.

271. A professor from the American University in Beirut, the plaintiff-abductee who was kidnapped and held for sixty-five days by members of an Iranian-sponsored terrorist organization, Hezbollah, suffered the same difficulties in distinguishing dream from reality. Regier v. Islamic Republic of Iran, 281 F. Supp. 2d 87, 91–92 (D.C. Cir. 2003). Unfortunately, when he was in the midst of a happy dream, he would wake up to a living nightmare. During the period of torture and imprisonment, Professor Regier testified to a “complete reversal from normal experience” with regard to the sleeping and the waking state. Id. at 92 (internal quotations omitted). “I would sometimes have a happy dream, and then I would wake up to the nightmare rather than the other way around. The nightmare was waking up, not going to sleep.” Id. (internal quotations omitted).

273. Id.
274. Id.
275. Id.
276. Id.
277. Id.
278. Id.
279. Id.
280. Id.

281. One wonders why the police allowed the investigation to go on for so long. As Gregory L. Lasak, the executive assistant to the Queens District Attorney said, “The detectives were very anxious to find the body. . . . They had questions about his credibility, but he also had the opportunity and he was confessing, and it would have been foolhardy to let him go.” Jan Hoffman,
also signed a written confession, and then later made a grisly videotaped confession in which he admitted to raping the girl on the couch, sodomizing her and smothering her with one hand, and then snapping her neck. From his dream, “I realized I killed her,” Smith wrote, “I feel sorry for what I did.”

The missing teenage girl happened to be watching television in her hide-out in Brooklyn and saw a video clip of Glenville Smith, her confessed killer, being escorted out of the precinct in handcuffs. “But I ain’t even dead,” she protested. Once she learned that she was presumed dead, she decided to go home, called her mother, and took the J train back to Queens. Despite his dream, apparently Smith had never touched the girl, and was eventually charged with obstructing justice, a felony, and filing a false police report, a misdemeanor. He was also examined by psychiatrists at Queens Hospital Center and Kings County Hospital.

Who knows what motivated Glenville Smith to confess to a crime that he did not commit? Perhaps he just wanted to be in the limelight; perhaps he was just crazy; perhaps Glenville Smith’s dreamed-of-deeds were merely repressed desires, so vividly reenacted in sleep, that he was genuinely unable to distinguish them from reality. Smith’s confusion, if confined to meditating about whether his hands were being held out before his real face or his dreaming face, presents an interesting philosophic quandary. But when the confusion is about whether he raped and killed his girlfriend’s daughter who had run away from home, and now he has gone to the police and signed a statement to that effect,


282. McAlary, supra note 273, at 16.
283. Id. Smith was having a tough time. When he returned home from looking for the girl’s body the first time, he was beaten up by the girl’s mother, his girlfriend. Id. His girlfriend later made a tape admitting that it was she, and not the police, who had beaten up Smith. Id.
284. Id.
285. Id.
286. Id.
287. Id.
288. Id.
289. The New York Times suggested that Smith “was fascinated by law enforcement. He claimed he was a security guard, and led Queens officers to 13 locations where the body might be buried. Then the teen-ager showed up, very much alive.” Hoffman, supra note 281, at B4.
confessing the dream of his crime—that kind of confusion could get a man executed in New York state.\textsuperscript{290} Had the girlfriend’s daughter decided to stay away forever, Glenville Smith’s horrible dream talk about her fate might have been adopted as the truth. An ambitious, zealous prosecutor might have decided to introduce his wild dream of raping and killing her—a wild dream made suddenly real by his waking confession—into evidence at his murder trial.\textsuperscript{291} Maybe Glenville Smith, given his predilection for taking responsibility, might have pled guilty. For a crime arguably punishable by the death penalty, the state might have insisted on going to trial. Either way—imprisoned for many years, or executed—it would have been a harsh punishment for an innocent man who could not distinguish dream from reality.

Perhaps he was caught in Zhuang Zhou’s dilemma and found himself wondering: am I Glenville Smith, just a guy hanging out at his girlfriend’s house in Queens, dreaming that I am a vicious killer? Or am I a vicious killer, dreaming that I am Glenville Smith, just a guy hanging out at his girlfriend’s house in Queens? There is an inherent duality in those questions—between what we call the real and the unreal, the

\footnotesize
\begin{itemize}
  \item \textsuperscript{290} In 1995, Governor George E. Pataki approved the long-inactive New York death penalty legislation. James Dao, \textit{Death Penalty in New York Reinstated After 18 Years; Pataki Sees Justice Served}, \textsc{N.Y. Times}, Mar. 7, 1995, at A1. New York limits the imposition of the death sentence to only those defendants convicted of first-degree murder. \textsc{N.Y. Penal Law} § 60.06 (McKinney 1998). To be convicted of first-degree murder in New York, the murder must have been committed under one of twelve enumerated aggravating factors, including that the killing was done in “an especially cruel or wanton manner pursuant to a course of conduct intended to inflict and inflicting torture upon the victim prior to the victim’s death.” Id. § 125.27(1)(a). One journalist wrote about the problem of executing innocent defendants:

  It almost happened in Queens, where prosecutors suddenly realized they had an innocent man on the track to the death house. The suspect must have been guilty, he admitted the crime. Only the rape and murder was a complete fabrication. He was saved by a girl who turned on the television during the middle of a thoroughly reckless teenage week and took the subway home.

  The next time you hear politicians talking about the social curing power of lethal injection, remind them of Glenville Smith, an admitted killer who might even be lucky enough to outlive his murder victim.

  \textit{McAlary}, supra note 273, at 16.

  \item \textsuperscript{291} In \textit{State v. Mitchell}, 577 N.W.2d 481 (Minn. 1998), for example, a fifteen-year-old defendant in a murder trial stated that he did not remember the night of the murder, but continued to have nightmares about what happened. The psychologist stated that Mitchell “feels haunted by the victim, [and is] terrified to go to sleep for fear of having more nightmares and [demonstrates] frequent vigilance in that he hears the victim’s voice whispering to him. There should be no mistaking that Mr. Mitchell experiences guilt.” Id. at 487. While there was other evidence that clearly inculpated Mitchell, the psychologist was allowed to testify that despite Mitchell’s lack of memory about the murder, the dreams evidenced his experience of guilt. Id.
\end{itemize}
Wild Dreamers

waking and the dreaming state, with their respective spheres of operation.

Zhuang Zhou characterized the duality differently from Descartes. For Descartes, there was a clear distinction between the dream and the waking life; the latter was “real,” and the former an illusion, and the difficulty was in distinguishing one from the other. Zhuang Zhou accepted the duality of the waking and the sleeping state, but he was much less certain about which state of consciousness staked a claim to “reality.” In another passage, Zhuang Zhou wrote:

While we dream we do not know that we are dreaming, and in the middle of a dream interpret a dream within it; not until we wake do we know that we were dreaming. Only at the ultimate awakening shall we know that this is the ultimate dream. Yet fools think they are awake, so confident that they know what they are, princes, herdsmen, incorrigible! You and Confucius are both dreams, and I who call you a dream am also a dream.

Zhuang Zhou’s suggestion that the waking state may be as illusory as the dreaming state is one that we see echoed in Tibetan Buddhism. Before I relate my one last story about dreams and the law—about Steven Linscott’s wild dreams and the terrible price he paid for sharing them—I offer one more interlude about how Tibetan Buddhists regard reality and the world of dreams. This system of thought shatters our western assumptions about the dualities of the real and the unreal, the waking and the dreaming state. The meaning of the dream is entirely different without these dualities. There is no longer any need to account for the relationship between the reality of the waking, conscious world and the sleeping, unconscious world. Both are equally unreal.

B. Interlude: Tibetan Buddhism

Tibetan Buddhism is a belief system that takes dreaming very seriously. Tibetan Buddhism is a part of the Mahayana tradition.

292. DESCARTES, supra note 262, at 145.
293. SOURCES OF CHINESE TRADITION, supra note 269, at 102–03.
294. Id.
295. The Shakyamuni Buddha himself was conceived in a dream. The Legend of the Buddha Shakyamuni, in BUDDHIST SCRIPTURES 34, 35 (Edward Conze ed. & trans., 1959). His mother, Queen Maya, was married to Shiddodana, the king of the Shakhas.
These two tasted of love’s delights, and one day she conceived the fruit of her womb, but without any defilement, in the same way in which knowledge joined to trance bears fruit. Just
having absorbed most of its ideas and traditions from Indian Buddhism during the seventh through the thirteenth centuries. One reason for the unique quality of Tibetan Buddhism is due to its interaction with the indigenous shamanistic religion of Tibet known as Bön. The shamans of the Bön religion sought to placate and manipulate the capricious spirits of nature. Magic, divination, and shamanistic possession are attributes of the Bön religion that permeate Tibetan Buddhism, creating a unique form of Mahayana Buddhism.

before her conception she had a dream. A white king elephant seemed to enter her body, but without causing her any pain.

296. Buddhism is generally divided into two traditions. First, Theravada Buddhism believes that the Buddha was a purified being, a saint or arhat, who finally escaped this corporeal life and endless cycle of life and death upon reaching nirvana. See ROBERT A.F. THURMAN, ESSENTIAL TIBETAN BUDDHISM 2 (1995). Under Theravada Buddhism, there were only a few other prior Buddhas prior to the Shakyamuni, but there are no living Buddhas today, and will be none until the future Buddha, Maitrya, arrives many thousands of years from now. Id. The second tradition is known as Mahayana, sometimes translated as “greater vehicle,” and it teaches that there are an infinite number of Buddhas. Id.; JOHN POWERS, INTRODUCTION TO TIBETAN BUDDHISM 313 (1995). Tibetan Buddhists believe that Buddhas could possibly show up here and now. THURMAN, supra, at 2. Many of their great religious figures became perfect Buddhas during their lifetimes, such as Padma Sambhava and Atisha. Id. at 3–4. Tibetan Buddhism also emphasizes another characteristic of Mahayana Buddhism, belief in what Thurman calls “Enlightenment Heroes,” the Bodhisattvas, disciples who mediate between Buddhas and the human population. Id. at 2–3.


298. Bön was considered a system of animistic and shamanistic practices, although even early on, Bön had assimilated a lot of Buddhist elements. POWERS, supra note 296, at 432. There are a lot of historical connections between shamanism and Hindu-Buddhist yoga practices. In North Asian shamanism, many concepts of Hindu and Buddhist Tantra were assimilated. Larry G. Peters, The Tamang Shamanism of Nepal, in SHAMANISM: AN EXPANDED VIEW OF REALITY 161, 172 (Shirley Nicholson ed., 1987). Both Tantric yoga and shamanism involve the embodiment of gods. Id. at 173. The shaman becomes possessed by the spirits he masters, while the yogi identifies “with gods representative of universal forces.” The altered state of consciousness that accompanies shamanistic possession in North Asian traditions usually occurs to a young person for the first time spontaneously and involuntarily, unlike the North American shamanistic traditions in which the young person takes off on a deliberate “vision quest.” Id. at 163. After shamanic training, the initiate can control his trance states, producing them deliberately in order to engage in healing rituals or produce visionary states. Id. at 168. The goals of the shaman are different from that of the yogi’s state of samadhi in which the yogi seeks enlightenment and detachment from the sensory world. Id. at 174. The shaman’s trance is “outwardly oriented. It is not autonomous, but is directed toward the community and therefore serves as a medium of communication between the supernatural or nonordinary reality and the community of men.” Id.

299. The first successful transmission of Buddhism into Tibet came during the reign of Songtsen Gampo who lived circa 618–650 C.E. POWERS, supra note 296, at 126. Both of his wives were Buddhists, one from Nepal and one from China, and they are considered by Tibetan tradition to be incarnations of the Buddha Tara. (One is called the Green Tara; the other the White Tara.) See Tara, at http://www.khandro.net/deities_Tara1.htm (last visited Apr. 7, 2004). Besides the more formal Bön religion, there were many animistic Tibetan folk religions, primarily concerned with the
Wild Dreamers

The dualities that we assume in the West, between the real and the unreal, and between their respective spheres of operation, the waking and the dreaming states, are rejected by Tibetan Buddhism. Under the Doctrine of the Dream-State from medieval Kagyudpa texts, the sensuous experiences of the waking state and the dreaming state are equally illusory. The Doctrine of Dreams follows from the Doctrine of Maya. Maya is explained in this way: “Worlds and universes are mind-made; they are of the stuff of which dreams are shaped. It is their illusiveness which is māyā. Things, appearances, are what mind makes them to be. Apart from mind they have no existence.” In the West, we

propitiation of spirits and demons. Powers, supra note 296, at 432–33. When Buddhism entered Tibet, there was no effort to suppress belief in these indigenous forces. Id. Rather, it incorporated them, making many of the spirits protectors of the dharma. Id. See generally Philip & Marcia R. Lieberman, Brief Introduction to Basic Concepts of Tibetan Buddhism (1999), at http://www.brown.edu/Research/BuddhistTempleArt/textfiles/Buddhism.txt (providing an introduction to Mahayana Buddhism).

300. There are four main schools of Tibetan Buddhism: Nyingma, Kagyu, Sakya, and Geluk. Powers, supra note 296, at 313. The Dalai Lama is the leader of the Gelukpa sect, and is regarded by most Tibetans as “the greatest symbol of all that is best and most worth preserving in their culture.” Id. at 412. While the four schools differ, they also share much in common, such as “the importance of overcoming attachment to the phenomena of cyclic existence and the idea that it is necessary for trainees to develop an attitude of sincere renunciation.” Id. at 313. All four sects also adhere to monastic discipline, and share the same corpus of religious texts. Id.

The Kagyudpa sect is characterized by its lineage of great teachers, beginning with Tilopa (988–1069 C.E.), and continuing with a long line of renowned students and teachers, including Naropa, Marpa, and Milarepa. Id. at 346–49. The founder of the Kagyudpa sect was Gampopa (1079–1153), who founded the Kagyudpa monasteries and imposed monastic discipline and a strict educational curriculum. Thurman, supra note 296, at 30. The Kagyudpa were also responsible for the development of the institution of official reincarnations that would have a “profound effect on Tibetan civilization.” Id. at 31.


302. Id.

303. Id. at 163. Maya is a central concept in Hindu philosophy. It represents the tendency of our egocentric, false concepts to keep us from perceiving the underlying unity of existence and the ultimate reality of Brahman. David R. Kinsley, Hinduism: A Cultural Perspective 92 (2d ed. 1993).

Maya . . . prevents us from being able to perceive ourselves from the perspective of eternity and enables us to remain content with the limitations of finite existence. In this sense, maya is a kind of drug that prevents us from becoming sick or dizzy in our endless rounds of birth and rebirth and keeps us firmly bound to the world of appearance and desire-motivated actions. Id. at 93.

The function of yoga is to help its practitioner see through this illusory world: “When, by means of yoga, the microcosmic aspect of mind is swept clear of the mists and the mirages of conditioned being, it sees itself as the One, emancipated from all māyāvic delusions, from all concepts of
assume that the dreaming state is unreal and the waking state is real. Tibetan Buddhism regards both as unreal, since both depend on the “same order of objective and sensuous perceptions, those of the dream-world merely being internal and those of the waking-world external.”

Both are merely appearances; the content of the waking state, as well as the dreaming state, is illusory.

The tradition of dream yoga originated with the eleventh century Indian yogi, Naropa; his teachings were later transmitted to Tibetan Buddhism under the name of the Six Principles of Naropa. Dream yoga is a technique that dreamers use to exercise their will on the content of the dream:

At the outset, in the process of realizing it to be māyā, abandon all feeling of fear [or dread]; And if the dream be of fire, transform the fire into water, the antidote of fire. And if the dream be of minute objects, transform them into large objects; Or if the dream be of large objects, transform them into small objects: Thereby one comprehendeth the nature of dimensions. And if the dream be of a single thing, transform it into many things; Or if the dream be of many things, transform them into multiplicity and of dualism, from all the magical deceptions of Nature.” The Path of Knowledge, supra note 301, at 163.

304. The Path of Knowledge, supra note 301, at 165.

305. GIUSEPPE TUCCI, THE RELIGIONS OF TIBET 26 (Geoffrey Samuel trans., 1980) (1970). Naropa is associated with the Kagyud school of Tibetan Buddhism, which gives particular emphasis to Tantric exercises and to the practices of yoga. Id. at 35. Naropa was an Indian Siddha, and his teachings are best known through the writings of his most famous student, Marpa (1012–1096). Id. at 36. On the various schools, see supra note 300.

306. TUCCI, supra note 305, at 98.

307. This form of dream yoga is similar to what some have called “lucid dreaming.” For a general introduction to lucid dreaming, see STEPHEN LABERGE, LUCID DREAMING (Jeremy P. Tarcher, Inc. 1985). Generally a lucid dream is a dream in which the dreamer is actively aware that he or she is in fact dreaming. This awareness is separate from the dream, and the dreamer can manipulate the story and characters to create a desired situation. Only about fifty-eight percent of all people have had a lucid dream once in their lifetime, and twenty-one percent have a lucid dream once or more in a month. If the dreamer has done either Buddhist or transcendental meditation, the average goes up to once or more a week. See DALAI LAMA, SLEEPING, DREAMING, AND DYING, supra note 4, at 103–04. The Dalai Lama postulated that, “[m]aybe this can be seen as an indication that these people have a higher degree of mindfulness. During the dream state there clearly is a form of consciousness in which one may engage in certain types of spiritual practice.” Id. at 104.
Wild Dreamers

a single thing:  
[Thereby one comprehendeth the nature of plurality and 
of unity.] 308

The dreamer thus learns through this psychic experimentation that the 
content of dreams is a mirage. 309 Naropa describes various practices in 
dream yoga, including special breathing, sleeping in a certain prescribed 
fashion, praying to one's guru "that thou mayest be enabled to 
comprehend the dream-state," and performing various visualizations 
during both waking and dreaming states. 310 Another practice of dream 
yoga is to "wake up" in the middle of the dream state and to recognize 
that you are dreaming, and then begin to rewrite the script. 311 Another 
mode of dream yoga is to convince yourself while you are awake that 
you are really dreaming. 312

The ultimate goal of dream yoga is to achieve a Dreamless State, one 
that is free from the illusoriness of both the waking and sleeping 
conditions, the Desireless State that goes "beyond both Form and Non-
Form." 313 Or in other words:

308. *The Path of Knowledge*, supra note 301, at 221.

309. The relationship between the dream state and maya can be summarized as follows:
The Doctrine of Dreams . . . further illustrates the Doctrine of Merry. It shows that even as all 
sensuous experiences of the waking-state are illusory, equally illusory are all sensuous 
experiences of the dream-state; these two states forming the two poles of human 
consciousness. In other words, Nature as a whole is the Dream of the One Mind; and until man 
conquers Nature, and thereby transcends myyā, he will remain asleep, dreaming the Dream of 
Ignorance. Whether in this world, or in any after-death state, all 
sangsaric, or 
karmically, 
conditioned experiences, are but dreams.

*Id.* at 164.

310. *Id.* at 216. The breathing is described as "pot-shaped" breathing. *Id.* at 217. The yogin 
should sleep "on the right side, as a lion doth," and press the thumb and ring-finger of the right hand 
against the throat arteries, stopping the nostrils with the left hand, and letting the saliva collect in the 
throat. *Id.* at 218.

311. Naropa describes in great detail a visualization that must be performed within the dream 

itself, concentrating the mind on a dot of various colors located in certain places on the body. *Id.* 

During the day, the practitioner of dream yoga must "hold to the concept that all things are of the 

substance of dreams." *Id.* at 216.

312. A fourteenth century Tibetan monk described the effects of such a practice:

"With unwavering attention hold to (the idea) that (everything is) like a dream 

Regardless of whether you are walking, sitting, eating, moving about or conversing. 

When you are not separate from the idea that 

Whatever is present, whatever is being done, and whatever is thought about is all a dream, 

Then you train yourself in absolute non-subjectivity (by realizing that your dream) 

Has no truth-value, is but something flimsy, something ethereal, something evanescent, 

Something fleeting, something faint."

*Anderson*, supra note 267, at 137 (quoting Longchenpa, a fourteenth century sage).

313. *The Path of Knowledge*, supra note 301, at 220–21 n.4.
The whole purpose of the Doctrine of Dreams is to stimulate the yogin to arise from the Sleep of Delusion, from the Nightmare of Existence, to break the shackles in which मायाः thus has held him prisoner throughout the aeons, and so attain spiritual peace and joy of Freedom, even as did the Fully Awakened One, Gautama the Buddha.314

The yogin comes to realize that all things perceived by the senses in the waking state are also a mirage because both the waking and the dreaming state are samsaric.315 The yogin will eventually enjoy as vivid consciousness in the dream-state as in the waking-state; and in passing from one state to another experiences no break in the continuity of memory. Thereby the content of the dream-state is found to be quite the same as that of the waking-state, in that it is wholly phenomenal and, therefore, illusory.316

Tibetan Buddhists believe that sleep, and dream yoga, are essentially a “rehearsal” for the process of dying.317 In certain kinds of special dreams, the subtle self actually departs from the gross body;318 a similar process happens upon death when the dead person enters the bardo.319 The bardo means “going between two” and generally refers to the mental voyage that an individual’s consciousness takes between death and the successive rebirth into another bodily form.320 The journey through the bardo is similar to the dreaming process in that it is entirely the production of the mind.321 The visions that one sees in the bardo—

314. Id. at 166.
315. In Hindu and Buddhist thought, samsara is usually defined as the endless cycle of birth and death. Ashok Gandadean, Comparative Ontology and the Interpretation of ‘Karma,’ 6 INDIAN PHIL. Q. 203, 211 (1979).
316. The Path of Knowledge, supra note 301, at 216 n.1.
317. DALAI LAMA, SLEEPING, DREAMING, AND DYING, supra note 4, at 43.
318. The subtle body is distinguished from the gross, or purely physical body. It exists “[w]edged between our familiar material universe and the ultimate Reality.” GEORG FEUERSTEIN, TANTRA: THE PATH OF ECSTASY 139 (1998). The subtle body is created by our own energy field and exists simultaneously with the physical body; it is “the energetic mold of the physical body.” Id. at 144. Tantric medical practitioners can prevent physical disease and mental imbalance by removing “energetic blockages and correcting damage on the level of the subtle body.” Id. at 145.
319. POWERS, supra note 296, at 286.
321. There are actually six bardo states: the bardo of birth, the bardo of dreams, the bardo of samadhi-meditation, the bardo of the moment before death, the bardo of dharmata, and the bardo of becoming. GURU RINPOCHE, THE TIBETAN BOOK OF THE DEAD: THE GREAT LIBERATION THROUGH
the smells and flavors, the fantastic sights of peaceful and wrathful deities—are all created in the same way our minds create dreams; they are merely the projections of the percipient. After death, as the consciousness seeks another incarnation, the bardo of dharmata, loosely translated as the “essence of reality,” is experienced. The reader of *The Tibetan Book of the Dead* calms the dead person by reminding him that the bardo is self-created:

> Now when the bardo of dharmatā dawns upon me,  
> I will abandon all thoughts of fear and terror,  
> I will recognise whatever appears as my projection  
> and know it to be a vision of the bardo;  
> now that I have reached this crucial point  
> I will not fear the peaceful and wrathful ones, my own projections.

Thus, for the Tibetan Buddhists, the practitioner of dream yoga is preparing himself for the passage from one lifetime to another. In the bardo, as in his dream practice, the body and mind separate and the true essence of reality “will appear, pure and clear yet hard to discern, luminous and brilliant, with terrifying brightness, shimmering like a mirage on a plain in spring.” Again, the reader of *The Tibetan Book of the Dead* seeks to reassure the dead person: “Do not be afraid of it, do not be bewildered. This is the natural radiance of your own dharmatā, therefore recognise it.”

From the mundane to the sublime. But now, on to our ultimate wild dreamer.

IV. THE ULTIMATE WILD DREAMER WHO LEARNS TO KEEP HIS MOUTH SHUT THE HARD WAY

The story of my last dreamer, Steven Linscott, is Kafkaesque. Here was a young man who spent over three years in prison because of his wild dreams. He could have spent more.

---


324. *Id.* at 40.
325. *Id.* at 41.
326. *Id.*
On the morning of October 4, 1980, Karen Anne Phillips, a twenty-four-year-old nursing student and practitioner of Kriya yoga, was found dead in her studio apartment in Oak Park on Chicago’s west side. Her body was lying flat, face down, and naked, except for a nightgown pulled tightly around her neck. There were numerous wounds and abrasions around her face, head, and back. Vaginal smears indicated that she had some kind of intercourse within the preceding twenty-four-hour period, although there was no evidence that she had been raped or sexually abused. Her assailant seems to have been someone known to the victim; there was no evidence of forcible entry into the apartment. A tire iron, encrusted with blood and hair, was found outside.

An autopsy revealed that Karen Anne Phillips had been killed by a combination of beating and strangulation. No one saw the assailant enter, but the next-door neighbor heard pounding about 1:00 a.m. on Saturday morning. He went and knocked on her door; the noise had ceased, but no one came to the door. Then when he returned to his apartment, the pounding noise started up again, and then it stopped. Karen Anne Phillips must have come to the end of her ordeal. When they found the body, the “decedent’s fingers were pressed together to form a hand signal, which in the Kriya Yoga religion signifies that the person was accepting death and seeking peace.”

The day after the murder, the police canvassed the area residents to see whether anyone had any information about what happened. They stopped at the Linscotts’ home, a few houses away from Phillips’ apartment. Steven Linscott and his wife, Lois, had just begun working as house parents in a Christian halfway house in the high-crime area of Chicago.

328. Id. at 1313.
329. Id.
330. Id. at 1306, 1313.
331. Id. at 1315.
332. Id. at 1313.
333. Id.
334. Id. at 1312.
335. Id.
336. Id.
337. Id. at 1313.
338. Id.
339. Id.
Wild Dreamers

Oak Park. 340 The police asked whether either of the Linscotts had seen anything unusual, "no matter how silly it might seem." 341 At that point, Steven Linscott suddenly remembered a vivid dream that he had experienced around 1:00 a.m. the night before, just when Phillips was murdered. 342 He did not tell the police about his dream at that time. 343 Later, he told someone who worked in the building about the dream, and then his wife. 344 Both of them thought that Steven Linscott should go to the police. His wife, Lois, "thought that he should tell the police about his dream if he thought it might be helpful to the police." 345 Big mistake. 346

Who was Steven Linscott? Here is how he was described by the court:

Defendant, Steven Paul Linscott, 26 years old at the time of the occurrence, was a Bible student at Emmaus Bible College in Oak Park. He lived with his wife and two children a few houses from the decedent’s apartment. Previously, defendant attended the University of Maine for two years and then joined the United States Navy. There, he became a radioman, and after background investigations, he was granted a top secret clearance. After an honorable discharge in March 1979, the

340. Id.
341. Id.
342. Id.
343. Id.
344. Id. at 1314.
345. Id.
346. The over eagerness of wild dreamers to share their dreams with police seems to be a recurring theme in the reported cases. One case that is frequently cited for the admission of dream narration testimony is Drake v. State, 476 So. 2d 210 (Fla. Dist. Ct. App. 1985). Drake is a perfect example of how an accused ought to learn to keep his mouth shut. The defendant was prosecuted for the attempted second-degree murder of his wife, Nancy, aggravated battery (a hammer to her head), and armed robbery with a deadly weapon of the ticket sales from a church concert that she had collected as church secretary of the First Nazarene Church of Winter Haven. Id. at 210. The defendant was having an affair with another woman, presumably the motive for his crimes. Id. at 212–13. The defendant gave a voluntary statement to the police, and the investigator testified that the defendant said, "I wake up at the nighttime, I wake up in the morning, I hear Nancy’s voice, she’s hollering for me to, ‘Stop, stop Tom, no Tom, no Tom, please Tom.’" Id. at 212. The defendant later denied making the statements; he also had no memory of the alleged incident. Id. at 214. In the interview, it was not clear whether Nancy’s voice haunted him after the crime, although it was implied. The statements were admitted as relevant to the defendant’s consciousness of guilt. Id. at 215. Although they were characterized as dream evidence, it is not so clear to me that they qualified. It sounded more like an auditory hallucination, or perhaps an auditory memory of the encounter in which he hammered his wife’s head. There is no evidence that the words were part of any nocturnal dream. Either way, he would have served himself better not to mention the dream/hallucination.
Linscotts lived in Maine. That fall, they moved to Chicago . . . . There is no criminality in defendant’s history, and he had never been arrested before this occurrence.347

By all accounts, Steven Linscott was squeaky clean. Here is how Steven Linscott first described his dream to the police: A short blonde man, about 5’5” to 5’7” tall, was talking to a second person.348 The short blonde man was wearing brown trousers, and a terry-cloth shirt with two or three stripes across his chest. In the beginning, he was calm, but then his demeanor changed. Linscott woke up and tried to shake his dream, then fell asleep again, and the dream resumed. The short blonde man took out an object, and began to strike the victim until blood was “flying everywhere.”349 There were similarities between the dream and the crime, but there were discrepancies as well, the most notable being that Linscott thought the victim was black, when in fact she was white.350

Steven Linscott was not surprised by his violent dream. In his words, “It wasn’t an unusual dream under the circumstances,” noting that he was nervous about having his family live with former prison inmates in a halfway house, in a neighborhood plagued by drugs and prostitution.351 “I believed that God maybe would use [the dream] in some way,” Linscott said, “I did not ever suspect that it would be used by the cops against me . . . If I have a sin, it was being naïve.”352

347. Linscott, 511 N.E.2d at 1313.
348. Sharon Cohen, Dream Leads to a Prison Term Until DNA Evidence Frees Him, L.A. TIMES, Sept. 18, 1994, at A1. At first, Linscott could not determine whether the second person was a man or a woman, although in his written account taken at the request of the police, he identified the victim as female. Id.
349. Id.
350. People v. Linscott, 500 N.E.2d 420, 423 (Ill. 1986). Linscott had several occasions to relate the dream to the police, and each time he did so, there were slight changes in the narrative. For example, the assailant’s shirt became a white terry-cloth shirt with two red and purple stripes on the chest, and one on the sleeve. Linscott, 511 N.E.2d at 1314. Some things were added in a later narrative as well. For example, Linscott told the police two days later that the victim was somewhat intelligent and seemed to be a religious person; he also described the living room in some detail. Id.
351. Barbara Sullivan, End of a Bad Dream, CHI. TRIB., June 24, 1994, § 5, at 1, 1–2. Lois Linscott was nervous about the neighborhood as well:
   It was a high-crime area; we could hear people arguing at night. Across the street was a prostitution area where pimps would come by and drop their women off and then pick them up later at night. Here we were, in this house with ex-convicts, suspicious of everyone around us, with two little kids. There was a lot of anxiety. It sets you up for nightmarish dreams. Id. at 2 (internal quotations omitted).
Wild Dreamers

Linscott revealed his naïveté by showing up at the police station two days later. After hearing Linscott’s initial narration of his dream, one of the police officers noticed “similarities between defendant’s [Linscott’s] physical appearance and the physical characteristics of the attacker who appeared in defendant’s dream. Defendant has straight blond hair, a light complexion, and a somewhat husky physique. He is just a bit under 6 feet tall.” After concluding their interview, the police asked Linscott to cooperate with their investigation, and he agreed to do so. Two days later, the police telephoned him and asked him to go down to the police station to tell them his dream a second time. The court noted, “When defendant arrived at the police station, he was wearing a light, colored terry-cloth shirt with a light colored stripe on a dark colored sleeve.” Why would Linscott have taken the risk of dressing like the assailant in his dream? How could he have been so clueless about the police officers’ suspicion that he was the man who killed Karen Anne Phillips? I am not sure “naïve” quite captures the nature of his folly.

Over a period of several days, Steven Linscott met with a detective and police officers, voluntarily giving them more than six hours of tape-recorded statements about the dream. According to Linscott, the police repeatedly told him he was not a suspect, and he voluntarily gave the police samples from his blood, saliva, and hair. At the police station, the officers accused Steven Linscott of having committed the murder, but he was not arrested until November 25, almost two months later. He rejected an offer to plea bargain, insisting on his innocence,
and voluntarily underwent and passed a lie detector test, which was not admissible at trial. From time to time, Steven Linscott may have equivocated on some of the details of the dream narration, but about one thing he was consistently and vocally adamant: he had not murdered Karen Anne Phillips.

At trial, the prosecution in the Linscott case relied on three pieces of evidence: first, the defendant’s account of his dream recorded by the police, which so closely paralleled the crime that “it was proof defendant committed the murder”; second, the head and pubic hairs that were found in the victim’s apartment; and third, the results of blood-typing tests from the seminal material taken from the victim. At trial, the prosecutor made certain erroneous statements about the blood and hair comparisons that eventually were deemed by the Supreme Court of Illinois to be improper and the basis for the remand in 1991 for a new trial. With respect to the hair comparison evidence, the prosecutor claimed in his closing argument that hairs “matching” the defendant’s had been found at the crime scene. However, there was no testimony at trial to support that statement, and in fact, there was evidence from several experts that no such identification was possible. Similarly, the prosecutor distorted the connection between the defendant’s blood and semen that was found around and inside the victim’s body, arguing again that there was something akin to a match—when in fact there was not. All that the evidence had shown was that Linscott could not be eliminated as a suspect—not that he could be identified as the perpetrator of this heinous crime.

363. See supra note 62.
364. Linscott, 566 N.E.2d at 1357.
365. Id.
366. Id.
367. Id. at 1358–64.
368. Id. at 1358.
369. For a full discussion of the subtleties of the hair comparison analysis, see id. at 1358–60.
370. Id. at 1362.
371. No objections were made to the prosecutor’s comments at trial, so the appellate court, and later the Supreme Court of Illinois, invoked the plain error doctrine since the error was egregious enough and “so substantial as to deprive defendant of a fair trial.” Id. at 1357 (citing People v. Sanders, 457 N.E.2d 1241 (Ill. 1983)).
The jury was swayed by this misleading and highly prejudicial physical evidence, and by Steven Linscott’s wild dream. He was convicted of the murder of Karen Anne Phillips and sentenced to prison for forty years. Linscott was transferred to the Centralia Correctional Center in southern Illinois, and his family moved down state, his wife Lois taking a job as a nurse. She found it difficult to suddenly be a single mother of three small children, and to have to explain what had happened to their father. In her words, “For the longest time when the kids were little, I would tell them the police made a big mistake . . . and they thought Daddy did something to hurt someone . . . . I tried not to

372. Id. at 1356. Using dream evidence to bolster otherwise scant empirical evidence has been the prosecutors’ strategy in other cases. In a case similar to Linscott’s, another defendant got into trouble for telling his dreams that related to a crime. In an appeal of a district court’s denial of Ralph S. Hockett III’s pro se petition for a writ of habeas corpus, the defendant challenged his plea of guilty of a murder in an Indiana state court because he was inaccurately advised that the state had in its possession a pair of tennis shoes stained with the blood of the victim. Hockett v. Duckworth, 999 F.2d 1160, 1161–62 (7th Cir. 1993). Hockett’s attorneys testified that even if the shoes were not stained with the same blood type as the victim’s, the nature and amount of evidence against the petitioner was such that they recommended he plead guilty in order to escape the death penalty. Id. at 1170. Besides his presence at the victim’s home and the fact that items of the victim’s property were found in his home, other factors that constituted the “totality of the evidence” included statements by the defendant, in one of which he “admit[ted] to having some dream about being there at the time [of the murders].” Id. at 1169.

In People v. Johnson, No. G030689, 2003 WL 21499895 (Cal. Ct. App. June 30, 2003), the trial court denied a motion in limine seeking to exclude statements by the defendant’s sister about a conversation they had several weeks after the brutal murder of their parents in which the defendant talked about his dreams of killing them. Id. at *4. At trial, Stacey, the sister, testified that the defendant told her, “I think I killed them.” Id. On cross-examination, Stacey testified: “I think he remembered maybe being in their house, but from the dream, in the dream . . . he thought it was a dream.” Id. at *5. Defendant also offered testimony from his psychiatrist that he was taking Prozac to control his depression, and that Prozac would cause “disturbing, vivid dreams of events that are not real.” Id. The court distinguished the case from one in which a defendant had taken sodium amytal, id. at *5–6 (citing People v. Johnson, 109 Cal. Rptr. 118, 199–20 (Cal. Ct. App. 1973)), and from one in which the defendant was under hypnosis, id. at *6 (citing People v. Blair, 602 P.2d 738, 753–54 (Cal. 1979)). The court also distinguished the case from State v. Tyler, 840 P.2d 413 (Kan. 1992), see supra note 78, in that Tyler was about the admissibility of a statement to prove the defendant’s state of mind before the murders, and here the defendant’s statement was offered as an “admission of involvement in the murders after the fact.” Johnson, 2003 WL 21499895, at *6. The reliability and credibility of the defendant’s dream talk “affected the weight, rather than the admissibility, of the evidence. The court properly permitted the jury to determine whether defendant’s statements described a dream or a real event.” Id. at *6–7. Unlike the Linscott case, however, there was substantial circumstantial evidence to prove that the dreaming defendant had indeed killed his parents. See generally id.


375. Id.
ever make it sound like the police were bad people.” 376 At first, Linscott was angry: “The bitterness was tearing me apart . . . . I realized I can’t keep going on like this or I’m going to go in a black fugue.” 377 “I’m not going to be able to handle life at all. I decided that as much as I hate it, I’m going to have to forgive these people (the authorities) . . . . Once I did, life got much better.” 378

After three and a half years in prison, the Illinois appellate court ordered a new trial, and Linscott was released on bond. 379 In a series of appeals, the prosecutors were sharply criticized for misrepresenting the physical evidence against Linscott, but during the next seven years, Linscott and his family lived in limbo, waiting for the new trial, wondering whether he would have to return to prison. 380 Finally, after more sophisticated DNA testing of the evidence, the state dropped the charges against him in 1992. 381 Almost thirteen years after Linscott’s arrest, the prosecutor explained, “The scientific evidence does not completely exonerate the man, but it raises sufficient doubts.” 382 It was

376. See Cohen, supra note 348, at A1 (internal quotations omitted).

377. See id. It is interesting that Linscott used the expression “black fugue.” A “fugue” is defined as “a state of psychological amnesia during which a patient seems to behave in a conscious and rational way, although upon return to normal consciousness he cannot remember the period of time nor what he did during it; temporary flight from reality.” WEBSTER’S NEW UNIVERSAL UNABRIDGED DICTIONARY 739 (2d ed. 1986). A “dissociative fugue” is one of the five types of dissociative orders (defined as a “break in the connection between memories or personality”). See Mary Eileen Crego, One Crime, Many Convicted: Dissociative Identity Disorder and the Exclusion of Expert Testimony in State v. Greene, 75 WASH. L. REV. 911, 913 n.16 (2000). The other four dissociative disorders are dissociative amnesia, dissociative identity disorder, depersonalization disorder, and dissociative disorder not otherwise specified. Id.

378. See Cohen, supra note 348, at A1 (internal quotations omitted).

379. There are a number of Illinois appellate decisions regarding the Linscott prosecution. In 1985, Linscott appealed his conviction by the Circuit Court of Cook County to the Appellate Court of Illinois, First District, Third Division, on the ground that the state had not proven him guilty beyond a reasonable doubt. People v. Linscott, 482 N.E.2d 403, 404 (Ill. App. Ct. 1985). His conviction was reversed, with one judge dissenting. Id. The Illinois Supreme Court granted the state’s petition for leave to appeal and reversed the appellate court’s decision. People v. Linscott, 500 N.E.2d 420, 421, 424 (Ill. 1986). The Illinois Supreme Court remanded the case to the appellate court, asking it to consider issues raised by the defendant that had not been previously addressed. Id. In the second round, the appellate court reversed and remanded for a new trial with directions. People v. Linscott, 511 N.E.2d 1303, 1304 (Ill. App. Ct. 1987). The Illinois Supreme Court granted leave to appeal, vacated the appellate court judgment, reversed the circuit court judgment, and remanded the case. People v. Linscott, 566 N.E.2d 1355, 1356, 1364 (Ill. 1991). In all, it took twelve years from the 1980 arrest for Linscott to be exonerated when the state finally abandoned the prosecution in 1992. Sullivan, supra note 351, at 2.


twelve years of hell to pay for having told the wrong people his wild dream.

How could this have happened? Because all of the attorneys involved in the Linscott case were rational empiricists. Justice McNamara’s dissent in the first appellate decision sheds some light: he unabashedly agreed with the state that Linscott’s dream narration constituted a confession, clearly rejecting the notion that Linscott could have obtained “knowledge” of the crime in any other fashion but through direct, sensory experience. Justice McNamara wrote:

I agree with the State that the defendant’s statements to the police were a confession. His statements, containing specific details of the crime, a psychological profile of the killer and an expression of concern about his hairs being found in the apartment, were not a recitation of a dream. The jury had the opportunity to observe the defendant, to hear his testimony, and to evaluate his testimony in light of all the other evidence. Their determination that the defendant acquired information about the murder, not through a dream, but because he was present, is very reasonable and understandable.

Neither could Justice Rizzi, the author of the majority opinion in both appellate decisions, bring himself to believe that Linscott could actually have dreamed about the details of a crime he did not commit. He suggested another way that Linscott might have gleaned the information: by reading about the crime in the newspaper. Justice Rizzi compared the language and details from various newspaper articles that had been published before Linscott went to the police with Linscott’s dream narration to demonstrate that Linscott was just reciting facts that were already publicly known. While Justices McNamara and Rizzi agreed that Linscott could only claim knowledge of the details of the crime on the basis of sense data, the nature of the sense data differed for each. For the dissent, the data was gathered from the scene of the crime itself—

383. *Linscott*, 482 N.E.2d at 408 (McNamara, J., dissenting).
384. *Id.* (McNamara, J., dissenting).
385. *Id.* at 406–07.
386. *Id.* at 406.
387. *Id.* The newspaper articles described the victim as a woman living alone, working as a nurse’s assistant. *Id.* The term “bludgeoned” to death was used; Linscott also used that expression. *Id.* Also, the articles indicated that the victim was “struck by a blunt instrument and that she suffered a head wound.” *Id.* In addition, the articles stated that “the police suspected that the victim knew her attacker because there were no signs of forced entry into the apartment.” *Id.*
and, therefore, tended to prove that Linscott was the likely perpetrator.\textsuperscript{388} For the majority, the data was gathered from the newspaper articles describing the crime—and, therefore, tended to prove only that Linscott was a reader.\textsuperscript{389}

I still do not know what to do with this story in my mind. I cannot make peace with it, confounded by the injustice of how things turned out in the “real world” for Steven Linscott. How could someone be sent to prison for murder on the basis of inexact DNA evidence and a wild dream? It shakes my faith in our criminal justice system and its alleged allegiance to rationality and an empirical episteme.

There was egregious misconduct on the part of the state’s attorney about the strength of the DNA evidence—the only empirical evidence introduced at trial. The only other evidence that the prosecution had in its arsenal was the dream talk. Perhaps the state’s attorney felt that he needed the dream talk because the physical evidence was so weak. Perhaps he was banking on at least some of the jurors having a premodern mindset; they could ignore the weakness of the DNA evidence based on the strength of the unknown and unknowable out-of-court declarant’s statement—the author of Steven Linscott’s dream. A number of other theories about dreams could have landed Steven Linscott in prison as well: for example, some Freudian version, that his dream was evidence of a Simpson-like “fatal obsession,” or that it expressed a wish that Karen Anne Phillips were dead. The use of this dream talk, along with the state’s attorney coming close to lying about how the hair and blood samples from the dead body of Karen Anne Phillips “matched” those of Steven Linscott, make his conviction particularly painful. The fact that the state’s attorney’s misconduct eventually resulted in the overturning of Linscott’s conviction, twelve years later, is of little comfort to me—and surely not to Steven Linscott and his family.

So even if—particularly if—I stay inside the dominant empirical episteme and the law’s commitment to rationality, what happened to Steven Linscott was a travesty of justice. To make the system work, we all have to play the game by the rules. The lawyers have to advocate with integrity. It would have helped if the defense attorney had been awake during the state’s closing argument to make an objection to the judge, who also must have been asleep, about what the prosecutor was

\textsuperscript{388} Id. at 408 (McNamara, J., dissenting).
\textsuperscript{389} Id. at 406–07.
Wild Dreamers

doing. If we are going to agree to only convict those defendants who have sufficient empirical evidence against them, then the state has to come up with it. If that empirical evidence is not there, then the defendant goes home. That is the deal we all signed on for.

Putting that aside—something I find extremely difficult to do—the wild dreamer in me wonders what exactly happened to Steven Linscott that October night over twenty years ago. If you throw open all the windows of your mind, there are many different theories about dreams that might apply. Perhaps Linscott was merely having a Freudian wish-fulfillment dream and expressing his unconscious desire that Karen Anne Phillips meet a violent end. Or is it possible that somehow the images of violence that took place that night passed through the walls of the buildings on her block and seeped into the sleep of Steven Linscott? Maybe Karen Anne Phillips herself sent him the film footage—her sheer terror projected the clip out into the universe, and he was the first receptive member of the audience to view it. Or is it possible that Steven Linscott ended up dreaming someone else’s dream—the dream of the perpetrator, the man whose hair and blood “match” the DNA found at the scene of the crime? Perhaps Steven Linscott was merely participating in the Jungian collective unconscious, dreaming the dream that belongs to us all? Perhaps he was dreaming a dream that a seventeenth century Iroquois might have considered a warning of impending violence. A Tibetan Buddhist might shrug his shoulders, and say: nothing happened to Karen Anne Phillips, either in her apartment that night or in Linscott’s dream—both realities, the waking and the dreaming, are equally illusory. Karen Anne Phillips does not exist, neither does Steven Linscott, and neither do you.

Steven Linscott himself had some views about the source of his dream that could hardly be described as mainstream. “We felt maybe God had provided this dream, where I could give a description,” Linscott commented in an interview, “Now I think it was demonic.”

390. See supra Part I.C.
391. See supra Part II.D.
392. See supra Part II.E.
393. See supra Part III.B.
394. Sullivan, supra note 351, at 1; see also supra Part I.A. Another defendant in a case involving dream evidence probably felt that “his” dream had a demonic source, although perhaps for a different reason. See Cruz v. County of DuPage, No. 96 C 7170, 1997 U.S. Dist. LEXIS 14397 (N.D. Ill. Sept. 15, 1997). In 1983, a ten-year-old child was abducted, sexually assaulted, and murdered, and grand jury indictments were obtained against the defendant, Cruz, and two others. Id. at *2. On appeal, Cruz alleged a number of things, including that the DuPage County prosecutors
Or there is another possibility that would satisfy even a radical empiricist: maybe at that time, Steven Linscott had a multiple personality.\textsuperscript{395} Maybe the second personality had an affair with Karen Anne Phillips, and for one reason or another, he killed her. Thus, what the first personality—the host—experienced with all the detail and vividness as a dream was indeed a memory, as the police, the prosecution, and the jury suspected; but not the memory of his own terrible experience—rather the memory of the other personality living within him.\textsuperscript{396} The two men lived side by side, immured from one another, the presence of the second personality being unknown to the

relied upon false evidence in deciding to go to trial. \textit{Id.} at \textsuperscript{395} *2–*3. Cruz’s conviction and death sentence were reversed by the Illinois Supreme Court not once, but twice, and at his third trial in 1995, Cruz was acquitted. \textit{Id.} at \textsuperscript{396} *3. In his Section 1983 action against DuPage County, Cruz alleged that prosecutors “knowingly participated in the fabrication of false evidence to inculpate Cruz,” including the fabrication of an “inculpatory statement by Cruz in the form of a ‘dream’” that Cruz had reportedly related to the police on May 9, 1983. \textit{Id.} at \textsuperscript{396} *4–*5. Cruz alleged that defendant Knight, a prosecutor, participated with the other defendants in the fabrication of this “dream statement,” through at least October 1995, and that he falsely reported to other DuPage County prosecutors that he had information that corroborated the “dream statement.” \textit{Id.} at \textsuperscript{396} *5. Knight’s motion to dismiss was denied. \textit{Id.} at \textsuperscript{396} *29. If Cruz’s version of the story is correct, here was a man who could have been put to death for a dream that he did not have at all—a dream that was fabricated by the state.

\textsuperscript{395} Multiple personality is the common name for what is called by mental health care professionals Dissociative Identity Disorder (DID). Crego, \textit{supra} note 377, at 912. A person suffering from DID may have from two to hundreds of personalities inhabiting one body; they may differ in age, race, and gender. \textit{Id.} at 913–14. The entire person, including all the personalities, is referred to as the “multiple,” and the “host” personality “refers to the personality in control of the body the greatest amount of time.” \textit{Id.} at 914. The other personalities are called “alters.” \textit{Id.} at 914–15.

\textsuperscript{396} The case of Billy Milligan, who was claimed to have twenty-four separate personalities, was one of the rare examples of a successful insanity defense for having a multiple personality. Juliette K. Orr, \textit{Comment, Multiple Personality Disorder and the Criminal Court: A New Approach}, 28 \textit{Sw. U. L. Rev.} 651, 657 (1999). Charged with raping three women near Ohio State University, Milligan was found innocent by reason of insanity. \textit{Id.} at 659. There was considerable public outrage at the verdict, which contributed to making it more “difficult for a multiple personality defendant to prevail with this defense.” \textit{Id.}

In a book written about the \textit{Milligan} case, one psychiatrist was interviewing “Billy,” when he was in high school, and asked him to describe how it felt to experience the trance-like periods where he could not really remember what happened. Daniel Keyes, \textit{The Minds of Billy Milligan} 160 (1981). “Billy” described it as being in a dream. \textit{Id.} When asked, “What are your feelings?” he responded:

Like a dream that comes and goes. My dad hates me. I hear him screaming. I have a red light in my room. I see a garden and a road—flowers, water, trees and nobody there to yell at me. I see lots of things that aren’t real. There’s a door with all the locks on it, and someone is pounding to get out. I see a woman falling down, and suddenly she’s turning into a pile of metal and I can’t reach her. Hey, I’m the only kid who can take a trip without LSD.

\textit{Id.}
first. Guilt and shame may have effaced the wall of amnesia between them; or perhaps it was a malevolent act of will on the part of the second personality, the perpetrator, who wanted to get the first Steven Linscott into trouble. Thus, Steven Linscott, the one who stood trial for the crime—the squeaky clean Bible student and father of three small children, the man who purportedly had a dream about the killing—was not in fact the person who committed the crime, but someone else who lived in his body. 397

We will probably never know who killed Karen Anne Phillips. I am willing to believe Steven Linscott that he did not. 398 I am willing to mourn for him, and with him, for the years he went to prison, and for the years those multiple appeals cast a dark shadow over his future. I am also resigned to live in a state of ignorance about how it happened that Steven Linscott had a dream about Karen Anne Phillips’ murder. I do not think the answer to that question is currently knowable, but I do believe there is an answer. Maybe our current understanding of human consciousness is too crude and too shallow to explain it now. Someday,

397. Somewhat related to an alter personality who commits the crime is when an individual who is in a deep sleep commits the crime. Known as the REM Sleep Disorder, it is possible for the normal muscle relaxation to fail during the dream stage of sleep, and the sleeper then acts out his dream. State v. Jones, 527 S.E.2d 700, 703 (N.C. Ct. App. 2000). In Jones, the defendant woke up to a loud bang and found a gun lying next to the face of his bleeding wife. Id. at 702. Jones said he did not remember shooting her, and a doctor testified that he suffered from REM Sleep Disorder, and that the sleeper did not always remember his dreams. Id. at 702–03. The defendant suffered from REM Sleep Disorder from two to three times a year, and had while sleeping, “kicked and damaged a wall, kicked a bedpost, squeezed and grabbed his wife and put his hand in her mouth, jumped out of bed and ran into a wall, and beat and scratched himself.” Id. at 703.

The jury was instructed on the defense of unconsciousness or automatism, but found the defendant guilty. Id. at 703–04. Under the defense of automatism in North Carolina, “when a person commits an act without being conscious of it, the act is not a criminal act even though it would be a crime if it had been committed by a person who was conscious.” Id. at 706; see also United States v. Blaney, 50 M.J. 533, 544 (A.F. Ct. Crim. App. 1999) (where defendant experienced a sleep-walking episode in which he engaged in oral sodomy in his sleep); People v. Saunders, No. 194742, 1998 WL 1991719, at *1 (Mich. Ct. App. Apr. 21, 1998) (where defendant acted out a sexual dream because of his REM behavior disorder and sexually abused his son); Hempel v. Palmateer, 66 P.3d 513, 513 (Or. Ct. App. 2003) (where defendant argued that if he touched the victim, he did so only in his sleep, while having a dream); State v. Taylor, No. M1999-02358-CCA-R3-CD, 2002 WL 799695, at *5–*6 (Tenn. Crim. App. Apr. 29, 2002) (where defendant argued she was having a night terror about stabbing an assailant when she was in fact stabbing her mother); supra note 2.

398. I know that the law does not agree with me, but if I were correct and Steven Linscott had a multiple personality, I would argue that it was not Steven Linscott who committed the crime, but one of his alternative, submerged personalities. If Linscott as host personality had no knowledge of the commission of the crime, it is contrary to our system of justice to hold him criminally liable for it. The most that the law would be willing to do with Dissociative Identity Disorder is to consider it as the basis for an insanity defense. See Orr, supra note 396, at 651.
some future historian will be reading through the judicial opinions and legal literature of the late nineteenth to the whatever centuries, looking for how our theory about consciousness manifested itself in judicial decisions, and will stumble across these meditations, and smile.

To my colleague and fellow time traveler who is not yet born, I salute you. We have all moved on by now, Steven Linscott, O.J. Simpson, Robert H. and Terry G., Ayla, Glenville Smith, the writer of this article, and the reader of this article—all the wild dreamers. You may find our questions quaint, but my guess is: you will not. You will know more than we do now about dreaming, but you will have your own state of ignorance to be resigned to. Almost four hundred years ago, Shakespeare wrote, “We are such stuff/As dreams are made on, and our little life/Is rounded with a sleep.”

Four hundred years from now, will you have figured out that sleep? The Tibetan Buddhists think there is a connection between our little life here—both in our waking and dreaming states—and that sleep. Have you figured that one out yet?

CONCLUSION

The sad truth is: wild dreamers do not fare all that well with the law. They would be well advised to keep their mouths shut. In all reported cases having to do with the admissibility of dreams, dream talk as evidence never inured to the benefit of any wild dreamer—whether it be in a criminal case or in family court. If the dream foretold of violence, the wild dreamer was always implicated in that violence—whether it happened or not. Even if his dream was prophetic of nothing in particular, the wild dreamer was punished for his prescience.

In short, all evidence of dream talk should be excluded from evidence. The law is too wedded to rationality; it belongs too squarely within the dominant empirical episteme. The law cannot accommodate any claim of knowledge considered “supernatural,” “archetypal,” or from the “unconscious.” Judges who live with, and inside, the law, are more than likely going to apply a theory of dreaming consistent with empiricism. If the dream predicts the future, the judge will probably opt for some version of a Freudian explanation and regard the dream as an expression of a wish. If the dream relates the past, the judge will probably regard the dream as a memory—in short, a confession.

399. WILLIAM SHAKESPEARE, THE TEMPEST act 4, sc. 1.
400. See supra Part I.C.
401. See supra Part IV.
To make matters worse, the members of the jury may ascribe to their own theories of dreaming, and apply them to the wild dreamer’s detriment. At least when judges are admitting dream evidence, their theory of dreaming is embedded in the record. Their justification for listening to a wild dreamer is subject to public scrutiny. Not so with the jury. They may draw all kinds of inferences about the source of the dream and its consequent meaning. They may be operating out of an entirely different set of epistemological and metaphysical assumptions. But we will never know because the jury operates secretly, and there is no need for them to justify their conclusions.

Criminal trials, as well as matters in family court, have a tremendous impact on the lives of the parties involved. In those cases, only the most reliable evidence should be admitted. Because everyone in the courtroom is sufficiently indoctrinated in the dominant episteme, we accept the requirement that the parties prove their case on the basis of sense data—on physical evidence, or on first-hand information about events observed. In this realm, there is some agreement about what constitutes a valid claim of knowledge, and what does not.

402. Jurors may also base their opinions about guilt or innocence on an incorrect understanding of one theory about dreaming or another. This is particularly true with some of the more difficult theorists such as Freud or Jung. See supra Parts I.C, II.D.


404. Our ability to perceive reality is particularly impaired during the transition between the sleeping and the waking states. Many cases involving dreams are about experiences related or utterances made just as the victim of a crime is either in, or coming out of, a dream state. The dream state has been used by the victim to explain why he did not protest some act of unwanted sexual aggression. See, e.g., United States v. Feltham, 58 M.J. 470, 471–72 (C.A.A.F. 2003) (where victim made inculpatory statements and remembered waking up from a wet dream at the same time as he was ejaculating into the defendant’s mouth); People v. Telle, 2003 WL 22922961, at *1 (Cal. Ct. App. 2003) (where victim did not protest a sexual assault because she was drifting “in and out of sleep, ‘felt like [she] was in a dream’ and wondered, ‘what is going on, why is this, who is this’”) (alteration in original); State v. Judge, 758 So. 2d 313, 316 (La. 1999) (where victim was asked repeatedly about her state of awareness because she was in a deep sleep when her uncle started kissing her); State v. Stevens, 53 P.3d 356, 359 (Mont. 2002) (where victim did not protest a sexual assault during a massage because she became “glued to the table” in a “deeply relaxed ‘dream state’ or ‘sleep rem stage’”); State v. Cooke, 785 A.2d 934, 937 (N.J. 2001) (where victim went back to sleep after an act of fellatio was performed on him during sleep because he thought “he may have had a bad dream”).

405. Suzanna Sherry made a similar point about the virtues of relying upon the rationalist/empiricist epistemology in public forum:

First, the question I am addressing is not whether the Enlightenment reliance on reason and empiricism is in fact the only epistemology, but whether we ought to proceed as if it were, at least in the public arena. Even if there are multiple and contradictory truths, some may be better suited for public adoption than others. There is a difference between objective truth and
Because there are a myriad of theories about the meanings of dreams, however, we have developed no criteria by which to judge their credibility. We share no assumptions about dreams—where they come from or what they mean—that are necessary to make sense of each other’s dream talk. Neither is there a dreaming culture in which to publicly debate and arrive at some consensus about the meaning of dreams. In Wittgensteinian parlance, we are not engaged—and cannot presently engage—in a language game about dreams. We have no shared store of symbols, the meaning of which is sufficiently determinate, to play that game. Even in therapeutic circles, the meanings of dream symbols are tossed about in what can best be described as a fluid professional debate.

And so I conclude: Given the dominance of rationalism and the empirical episteme in our courts of law, dream talk should not be admissible. In modern western culture in the early part of the twenty-first century, we are just not ready to make dreams the basis of knowledge, at least not in a public arena where the task is the resolution of conflict and the meting out of justice. Jurors are too unpredictable. We do not know what their theories about the meaning of dreams might be, and their silent role at trial allows for no way to detect their metaphysical and epistemological assumptions. We therefore cannot allow the determination of guilt or innocence to be based upon dream talk. Wild dreamers abound. There is too much at stake.

It is unlikely that the situation will change for wild dreamers in the near future. I do not think the dominant episteme is going to soon yield its firm grasp on the imagination of most of those who live with and in the law. Neither do I think it is going to soon yield its firm grasp on the imagination of almost anyone else who belongs to the dominant culture. On the other hand, empiricism is not here to stay forever—not if you take seriously the history of ideas. The history of ideas suggests that our beliefs about ways of knowing, and about how the world is put together and why, will continue to change over time.407 My guess is that we are in

---

406. See LUDWIG J. WITTGENSTEIN, PHILOSOPHICAL INVESTIGATIONS § 2, at 3e (1953).
407. The traditional view of the evolution of science was that its progress was gradual and linear. Thomas Kuhn turned that traditional view on its head in his work, The Structure of Scientific Revolutions. THOMAS S. KUHN, THE STRUCTURE OF SCIENTIFIC REVOLUTIONS (3d ed. 1962); see Dudley Shapere, The Structure of Scientific Revolutions, 73 PHIL. REV. 383 (1964) (discussing Kuhn). According to Kuhn, who based his theory on the hard sciences such as physics and
Wild Dreamers

the halcyon days of the empiricist regime, but that within a century or two, perhaps sooner, we might well discover that our original premise—that there are only five senses with which to collect data—is inadequate. We might well discover that the union of our human body and mind and spirit has far more potential as a percipient being than we have ever imagined, and that from an enhanced understanding of our consciousness, we will also discover new ways of knowing.

But we are incarnate, and captives of our own particularity. Each of us has been born into a certain body that begins and ends its life over a certain historical period that dictates how we experience the world and give it meaning. Lawyers practicing in the early twenty-first century are no exception. It is our job to understand how the law in our historical period is defined by the dominant culture’s episteme and metaphysics, and how our legal system operates. It is our job to help clients use that legal system, to settle their disputes, to distribute wealth, to determine who raises their children, to remedy wrongs, and when harm is done, to ensure that justice prevails. For those lawyers who try cases, that means they must know the rules of evidence and apply them with skill as they zealously represent their clients.

And here is where the role of the wild dreamer comes in, just in case you had despaired. We do have utility. A lawyer armed with the knowledge that there are alternative epistemes and competing metaphysical assumptions, is a better lawyer. For one thing, when dream talk comes up in a case, the wild dreamer will already be acutely aware of the potential harm that a cleverly placed dream narration might do to the client—and will loudly protest its introduction into evidence. And if in need of authority, these meditations can serve as footnote fodder. O.J. Simpson’s Dream Team had almost nothing to rely upon. See supra notes 66–70 and accompanying text.

chemistry, change in science does not always follow a logical, gradual progression, but is marked by abrupt, almost revolutionary, changes in paradigms. Shapere, supra, at 391–93. The dominant paradigm, or widely accepted model of looking at reality, with its shared assumptions that best describe the world, and how to learn more about it, are challenged by competing schools, causing an eventual paradigm crisis. Id. Eventually, the dominant paradigm is thrown out, and the scientific community embraces a radically different world view. Id.
immediately what to advise the client once the specter of litigation or prosecution loomed: Shut up.

In the process of writing this article, I have also answered my own submerged questions about the place of wild dreamers in the academy. We have three sacred tasks. Count them: one, two, three; I am a lawyer after all. I know that you do not trust me—maybe the one, two, three of things will make you feel safer.

The three sacred tasks of the wild dreamer law professor are as follows.

One: Academic wild dreamers have the sacred task of training other wild dreamers to practice law without losing their hearts and souls to the dominant episteme and its bleak metaphysical underpinnings. Sometimes the wildest of dreamers drift into law school, and find it a cold and lonely place; they quickly discover that their ways of experiencing and understanding the world are regarded not only as foreign, but unacceptable. They leave or drop out; sometimes they just give up on their alternative ways of knowing and sign on, part and parcel, to the dominant episteme. The profession is the worse for it. We need to nurture and protect them. Dullards who live with and in the law abound. Wild dreamers do not. Our profession needs the outsider’s perspective of the wild dreamer, as well as their propensity to ignore the sharp distinctions between established categories of thought—and to redefine them.

Two: Academic wild dreamers have the sacred task of engaging in dream talk under the guise of legal scholarship. We have the time and space both to daydream, and to dream in our sleep. Not only that, we are expected to travel through time in both directions—to figure out how and why things were done in the past, and to imagine how and why things might be done in the future. We frequently engage in

409. More wild dreamers are lured into the humanities or the arts; professions like the law and medicine do not usually call their names. James believes that the range of mystical experiences is much wider than most people believe, from the full blown revelatory experiences of the Christian saints and Hindu or Buddhist yogis, to the experiences that we have in our everyday lives. JAMES, supra note 5, at 295. He gives examples:

Most of us can remember the strangely moving power of passages in certain poems read when we were young, irrational doorways as they were through which the mystery of fact, the wildness and the pang of life, stole into our hearts and thrilled them. . . . We are alive or dead to the eternal inner message of the arts according as we have kept or lost this mystical susceptibility.

Id. There are some poets who enter law school, and many of them worry that the experience will beat out of them altogether their receptivity to the eternal inner message of the arts.
Wild Dreamers

conversations with the dead,\textsuperscript{410} and the not-yet-living. These are luxuries not afforded to the practitioner. The conjunction of freedom and time and imagination—that is our greatest gift, and our greatest responsibility. It enables those of us lucky enough to spend our lives in the academy a chance to drift endlessly in the dream world. Some may laugh and say: Dreamy drifting could hardly be described as socially useful. But they would be wrong. That dreamy drifting may take us out to uncharted seas—to find new routes for passage, escape, perhaps for salvation.

Three: There is no three. I just made up the list of three so that you would feel better. And perhaps to leave you with an empty space—a space for you to confront some empty space, perhaps to have a wild dream in.

\textsuperscript{410} After Machiavelli had just been released from prison, he wrote:

I return to my house and go into my study. At the door I take off my clothes of the day, covered with mud and mire, and I put on my regal and courtly garments; and decently reclothed, I enter the ancient courts of ancient men, where, received by them lovingly, I feed on the food that alone is mine and that I was born for. There I am not ashamed to speak with them and to ask them the reasons for their actions; and they in their humanity reply to me. And for the space of four hours I feel no boredom, I forget every pain, I do not fear poverty, death does not frighten me. I deliver myself entirely to them.

Letter from Niccolò Machiavelli to Francesco Vettori (Dec. 10, 1513), \textit{in NICCOLÒ MACHIAVELLI, THE PRINCE} 107–08 (Harvey C. Mansfield, Jr., trans., 1985). I found this quote in a wonderful piece by Mary Ann Glendon in which she extols the virtues (and vices) of interdisciplinary scholarship. She writes about finding this letter written by Machiavelli: “On the other hand, one need not find one’s intellectual companions in one’s own discipline, or one’s own nation-state, or even in one’s own time. In fact, the most moving account of reaching across margins that I have ever seen is about friendship with the dead.” Mary Ann Glendon, \textit{Why Cross Boundaries?}, 53 WASH. & LEE L. REV. 971, 979 (1996).
FOSTER V. CARSON: THE NINTH CIRCUIT MISAPPLIES THE CAPABLE-OF-REPETITION-YET-EVADING-REVIEW EXCEPTION TO THE MOOTNESS DOCTRINE AND LENDS A FREE HAND TO BUDGET-CUTTING STATE OFFICIALS

Joshua C. Gaul

Abstract: In Foster v. Carson, public defender organizations and indigent defendants sued the chief justice of the Oregon Supreme Court for suspending appointments of indigent defense counsel. Before the parties could fully litigate the case, the chief justice reinstated appointments. Subsequently, the United States Court of Appeals for the Ninth Circuit dismissed the case as moot and held that the exception to the mootness doctrine for cases capable-of-repetition-yet-evading-review did not apply. A case falls under that exception when the party resisting mootness demonstrates that it was not possible to fully litigate the action before it ceased and there is a reasonable expectation that the party will be subjected to the same action in the future. Because the court concluded that it was not possible to fully litigate the case before the chief justice reinstated appointments, applicability of the capable-of-repetition-yet-evading-review exception depended only on whether there was a “reasonable expectation” that the injury would recur. When evaluated in light of U.S. Supreme Court and Ninth Circuit precedent, the facts in Foster support a finding that there was a reasonable expectation that the chief justice would again suspend funding for indigent defense counsel. The public interest in deciding the constitutionality of the chief justice’s action further supports application of the exception.

Thanks to a prolonged economic recession, the escalating cost of domestic security, and unrelenting pressure not to raise taxes, state governments face crushing budget deficits. The states faced a collective $200 billion shortfall for fiscal years 2001 through 2003. As states struggle to find ways to reduce deficits, state programs and services are being reduced or altogether eliminated. State and local judicial systems are not immune to these cuts.

In response to Oregon’s severe budget crisis, the Oregon State Legislature held five special sessions during the second half of 2002 to

2. Broder, supra note 1, at A12.
3. See id.; Janofsky, supra note 1, at A20.
4. See Foster v. Carson, 347 F.3d 742, 744 (9th Cir. 2003).
address the budget shortfall. During these sessions, the legislature substantially cut funds available to the Oregon Judicial Department for its indigent defense counsel programs. As a result, the chief justice of the Oregon Supreme Court, acting in his capacity as the administrative head of the Oregon Judicial Department, suspended appointments of indigent defense counsel between March 1, 2003 and June 30, 2003. Indigent defendants charged with a variety of crimes were denied counsel, and their cases were delayed until after June 30.

Before the chief justice reinstated funding, a group of indigent defendants and defense counsel organizations had sued him in Foster v. Carson. The plaintiffs claimed that the chief justice had violated their constitutional rights to counsel, due process, and equal protection when he suspended appointments of indigent defense counsel. After the U.S. District Court for the District of Oregon had granted defendant’s motion to dismiss, plaintiffs appealed to the U.S. Court of Appeals for the Ninth Circuit. The Ninth Circuit dismissed the case as moot.

The U.S. Supreme Court recognizes a number of exceptions to mootness, including the exception for cases capable-of-repetition-yet-evading-review. For a case to fall under this exception, the party resisting mootness must demonstrate that the challenged action is too short in duration to be fully litigated before it ceases and that there is a reasonable expectation that the party will be subjected to the same action in the future. The Foster court held that the exception did not apply because there was no reasonable expectation that the plaintiffs would again face a suspension of indigent defense counsel appointments.

---

5. Id. at 748; see also Appellants’ Consolidated Opening Brief at 4–5, Foster (Nos. 03-035457 & 03-035458).
7. Foster, 347 F.3d at 744.
8. Id.; see also infra note 135.
9. Foster, 347 F.3d at 744. This Note does not address the constitutional issues raised by the Foster plaintiffs or the plaintiffs’ right to speedy trials.
10. 347 F.3d 742 (9th Cir. 2003). In addition to the chief justice, the plaintiffs also named the state court administrator and several Oregon judges as defendants. Id. at 742.
11. Id. at 745.
12. Id. at 749.
13. Id.
14. See infra Part I.A.
16. Foster, 347 F.3d at 748–49.
An Exception to Mootness

This Note argues that the Ninth Circuit erred in dismissing the *Foster* case as moot. Part I describes the mootness doctrine and details the capable-of-repetition-yet-evading-review exception as applied by the U.S. Supreme Court and the Ninth Circuit. Part II details the facts, procedural history, and holding of the *Foster* decision. Part II also summarizes information submitted to the *Foster* court by plaintiffs’ counsel about Oregon’s ongoing budget crisis. Part III then argues that under the reasonable expectation standard established by the U.S. Supreme Court and the Ninth Circuit, the issue in *Foster* is capable-of-repetition-yet-evading-review. Part III also argues that the public interest in settling the constitutionality of suspending indigent defense counsel appointments further weighed against holding the case moot. This Note concludes that the Ninth Circuit should have applied the capable-of-repetition-yet-evading-review exception instead of dismissing the case as moot.

I. CASES THAT ARE CAPABLE-OF-REPETITION-YET-EVADING-REVIEW ARE NOT MOOT

To satisfy the U.S. Constitution’s jurisdictional requirements, federal courts must find both that the parties bringing the action have standing and that the action either is not moot or fits within one of the exceptions to the mootness doctrine. 17 Standing addresses whether the plaintiff is the proper party to bring the case before a federal court for adjudication. 18 To determine whether a case is moot, a court asks whether the “issues presented are no longer ‘live’ or the parties lack a legally cognizable interest in the outcome.” 19 For example, a case will become moot if an essential party dies during the appeals process 20 or if a student challenging a school policy graduates from the school. 21

The U.S. Supreme Court has articulated four exceptions to the mootness doctrine. 22 One of these exceptions is that the action is

18. *ERWIN CHERERINSKY, CONSTITUTIONAL LAW PRINCIPLES AND POLICIES § 2.5.1, at 60 (2d ed. 2002).
22. *See CHERERINSKY, supra* note 18, §§ 2.7.2–.5, at 114–27.
capable-of-repetition-yet-evading-review. To establish that a case falls into the capable-of-repetition-yet-evading-review exception, the party resisting mootness must first demonstrate that it is not possible to fully litigate the challenged action before it ceases. Next, the party must show that there is a reasonable expectation that it will be subjected to the challenged action in the future. U.S. Supreme Court and Ninth Circuit cases applying the capable-of-repetition-yet-evading-review exception establish what constitutes a “reasonable expectation.” Federal courts do not recognize an exception to mootness for cases involving a strong public interest. However, both the U.S. Supreme Court and the Ninth Circuit have held that a strong public interest in settling the legality of an action may weigh against a holding of mootness.

A. U.S. Supreme Court Precedent Establishes a Flexible Doctrine of Mootness

Article III of the U.S. Constitution limits the jurisdiction of federal courts to “cases and controversies.” Cases and controversies consist of active disputes between two or more adversarial parties. Limiting courts to resolving only live disputes helps to ensure that the judicial branch does not intrude upon the legislative and executive branches of the federal government. To maintain their case or controversy, all parties must retain a personal stake in the litigation throughout all its stages.

Courts have developed the related doctrines of standing and mootness to determine whether the issue being litigated presents a case or

23. Id. § 2.7.3, at 117.
25. Id.
31. Id.
An Exception to Mootness

controversy. The doctrines’ common nexus in Article III has led the U.S. Supreme Court to describe mootness as “standing in a time frame.” The Court has stated that “[t]he requisite personal interest that must exist at the commencement of litigation (standing) must continue throughout its existence (mootness).” Although the Court has frequently emphasized the relationship between mootness and standing, it consistently treats mootness as a more flexible doctrine.

The U.S. Supreme Court has articulated four exceptions that allow a federal court to retain jurisdiction over an otherwise moot case. A case will not be moot when (1) the injured party faces collateral consequences stemming from the injury; (2) the case is certified as a class action; (3) one party voluntarily ceases the allegedly illegal action; or (4) the action at issue is capable-of-repetition-yet-evading-review.

33. See Friends of the Earth, Inc. v. Laidlaw Envtl. Servs. (TOC), Inc., 528 U.S. 167, 180 (2000). There is some debate over whether the mootness doctrine is mandated by Article III. In Honig v. Doe, 484 U.S. 305 (1988), Chief Justice Rehnquist argued that the mootness doctrine does not derive from Article III: it is a prudential rule with “an attenuated connection [to Article III] that may be overridden where there are strong reasons to override it.” Id. at 330–31 (Rehnquist, C.J., concurring). For a discussion of the constitutional roots of mootness, see generally Evan Tsen Lee, Deconstitutionalizing Justiciability: The Example of Mootness, 105 HARV. L. REV. 605 (1992).


35. Id. The definition of mootness as “standing in a time frame,” which has been widely adopted by the lower courts and legal scholars, has been questioned by the Court. In Friends of the Earth, Inc. v. Laidlaw Environmental Services (TOC), Inc., the Court suggested that the definition is not comprehensive. 528 U.S. at 190.

36. See Geraghy, 445 U.S. at 400–01; see also Roe v. Wade, 410 U.S. 113, 125 (1973). There are circumstances where the likelihood that a defendant will resume the challenged behavior is too slim to support standing, but would be sufficient to save a case on appeal from dismissal as moot. Friends of the Earth, 528 U.S. at 190.

37. See CHEMERINSKY, supra note 18, §§ 2.7.2–.5, at 114–27.

38. See, e.g., Spencer v. Kemna, 523 U.S. 1, 7–10 (1998) (outlining the collateral consequence exception for both civil and criminal cases).

39. See, e.g., Sosna v. Iowa, 419 U.S. 393, 401–02 (1975) (holding that a court retains jurisdiction over a class action suit even if the named party’s case has become moot).

40. See, e.g., Friends of the Earth, 528 U.S. at 189–90 (holding that a citizen suit under the Clean Water Act did not become moot when the permit holder began complying with its permit or when it shut down the offending facility).

41. See, e.g., Honig v. Doe, 484 U.S. 305, 318–20 (1988) (holding that a student’s challenge of his suspension was not moot because the student retained the option to re-enroll in a California school).
B. The Capable-of-Repetition-yet-Evading-Review Exception Requires There To Be a Reasonable Expectation That the Action Will Recur Rather Than a “Demonstrated Probability” of Recurrence

The U.S. Supreme Court has developed a two-part test to determine if a case falls into the capable-of-repetition-yet-evading-review exception. First, the duration of the action at issue must be too short for the parties to fully litigate its legality before it ceases. Second, there must be a reasonable expectation that the party challenging the action will be subjected to the same action again. Establishing such reasonable expectation does not require a party to provide a “demonstrated probability” of recurrence.

In Honig v. Doe, the Court concluded that establishing a reasonable expectation that an action will recur does not require a demonstrated probability of recurrence. The Honig Court accepted a chain of probable events leading to the recurrence of the challenged action as sufficient to establish a reasonable expectation. In Honig, two California students, John Doe and Jack Smith, sued the California superintendent of public instruction after they were expelled from high school for behavior related to their emotional disabilities. Both students alleged that their expulsions violated the Education of the Handicapped Act (EHA). After the district court and the Ninth Circuit had ruled in their favor, the U.S. Supreme Court granted certiorari. At oral argument, the United States, appearing as amicus curiae, argued that the case was moot because Doe was no longer entitled to protection under the EHA due to his age, and Smith had moved out of the school district from which he had been expelled.

43. Id.
44. Id. Because the Foster court accepted that the suspension of funds for defense counsel was too short to be fully litigated, this Note focuses only on the court’s application of the “reasonable expectation” requirement. Foster v. Carson, 347 F.3d 742, 746 (9th Cir. 2003).
45. Honig, 484 U.S. at 320 n.6.
47. Id. at 320 n.6.
48. See id. at 321–23.
49. Id. at 312–16.
50. Id.
51. Id. at 316–17.
52. Id. at 318 nn.5–6.
An Exception to Mootness

The Court held that the case was moot for plaintiff Doe.53 The EHA only applies to students between the ages of three and twenty-one, and Doe was twenty-four by the time the Honig case reached the Court.54 Because the EHA no longer applied to Doe, there was no possibility that California would again deprive him of his rights under the EHA.55

However, because plaintiff Smith was only twenty when Honig reached the Court,56 it was possible that California could again deny him his rights under the EHA.57 The Court held that the capable-of-repetition-yet-evading-review exception applied to Smith58 because there was a reasonable expectation that Smith would again be expelled from a California school.59 For that to happen, the Court assumed that Smith would again enroll in a California school (even though he had not stated that he would do so), be placed in an environment where his behavior could not be adequately controlled, behave in an aggressive manner caused by his disability, and be expelled for that behavior.60 Despite these assumptions, and the fact that Smith was nearly too old for the EHA to apply to him, the Court concluded that there was a reasonable expectation that a California school would again expel him.61

C. Under Ninth Circuit Precedent, Courts May Find That a Reasonable Expectation Exists When There Is Some Indication That the Challenged Action Will Be Repeated, but Not When Recurrence of the Action Is “Highly Unlikely” or “Speculative”

The Ninth Circuit has held that a party challenging an action must demonstrate “some indication that the challenged conduct will be repeated” in order to satisfy the reasonable expectation requirement of the capable-of-repetition-yet-evading-review exception.62 Like the U.S. Supreme Court, the Ninth Circuit does not require a demonstrated

53. Id. at 318.
54. Id.
55. Id.
56. Id.
57. See id. at 320–23.
58. Id. at 323.
59. Id. at 319–20.
60. See id. at 320–23.
61. Id. at 323.
62. Alaska Ctr. for the Env’t v. United States Forest Serv., 189 F.3d 851, 856 (9th Cir. 1999).
probability of recurrence. 63 The Ninth Circuit will not, however, apply the exception when repetition of the challenged action is “highly unlikely”64 or “too speculative to prevent mootness.”65 Determining whether a case has become moot often involves a “highly individualistic . . . appraisal of the facts.”66

The Ninth Circuit has established a low threshold for demonstrating a reasonable expectation that a challenged action will recur.67 That standard is met when the party challenging the action establishes “some indication” that the action will be repeated.68 For example, in Miller ex rel. NLRB v. California Pacific Medical Center,69 the court applied the U.S. Supreme Court’s holding in Honig that the complaining party “need not show that there is a ‘demonstrated probability’ that the dispute will recur.”70 In Miller, the National Labor Relations Board (NLRB) sought a preliminary injunction to restore the collective bargaining status of the California Nurses Association (CNA) with the California Pacific Medical Center (CPMC) pending resolution of an unfair labor charge.71 CNA lost its collective bargaining status during the merger of two other hospitals into CPMC.72 Prior to an en banc rehearing, the NLRB issued its final decision on the underlying unfair labor charge, which rendered the preliminary injunction unnecessary.73 Because CPMC continued to undergo restructuring and because it held contracts with five other unions, the Ninth Circuit concluded that there was a reasonable expectation that the NLRB would again seek injunctive relief against CPMC in an unfair labor proceeding.74

63. Barilla v. Ervin, 886 F.2d 1514, 1520 (9th Cir. 1989) (discussing Honig).
64. Unabom Trial Media Coalition v. United States Dist. Court for the E. Dist. of Cal., 183 F.3d 949, 953 (9th Cir. 1999); Sze v. INS, 153 F.3d 1005, 1009 (9th Cir. 1998).
65. Dilley v. Gunn, 64 F.3d 1365, 1369 (9th Cir. 1995).
66. Alaska Ctr., 189 F.3d at 856 (quoting 13A CHARLES ALAN WRIGHT & ARTHUR R. MILLER, FEDERAL PRACTICE AND PROCEDURE § 3533 (2d ed. 1984)).
67. See Barilla, 886 F.2d at 1520.
68. Alaska Ctr., 189 F.3d at 856.
69. 19 F.3d 449 (9th Cir. 1994) (en banc).
70. Id. at 454 (quoting Barilla, 886 F.2d at 1520 (discussing Honig v. Doe, 484 U.S 305 (1988))).
71. Id. at 451.
72. Id.
73. Id. at 453.
74. Id. at 454.
An Exception to Mootness

The Ninth Circuit will not apply the exception when a recurrence of the challenged action is “highly unlikely” or “speculative.” In *Mayfield v. Dalton*, the court noted that “speculative contingencies afford no basis for . . . passing on the substantive issues” of a case. The plaintiffs, two marines, challenged the constitutionality of a Department of Defense (DOD) policy requiring tissue samples for DNA analysis from all members of the armed services. By the time the case reached the Ninth Circuit, both plaintiffs had left active duty, allowing the DOD to argue that the case had become moot. The marines countered that their case was not moot because they were subject to recall to active duty in the event of a national emergency. The *Mayfield* court held that the case was moot because there was no reasonable expectation that the marines would again have to submit DNA samples. The marines’ recall to active duty “could happen only at some indefinite time in the future and then only upon the occurrence of future events now unforeseeable.” In addition, the Ninth Circuit concluded that several changes made by the DOD to its DNA sampling program after the plaintiffs had brought suit had “materially alter[ed] many aspects of the policy that [the marines] challenged” and rendered it unlikely that they would again be subjected to the challenged DNA sampling program.

Similarly, in *Dufresne v. Veneman*, the Ninth Circuit held that the case was moot after concluding that the recurrence of the challenged

75. Unabom Trial Media Coalition v. United States Dist. Court for the E. Dist. of Cal., 183 F.3d 949, 953 (9th Cir. 1999); Sze v. INS, 153 F.3d 1005, 1009 (9th Cir. 1998).
76. Dilley v. Gunn, 64 F.3d 1365, 1369 (9th Cir. 1995).
77. 109 F.3d 1423 (1997).
78. Id. at 1425 (quoting Preiser v. Newkirk, 422 U.S. 395, 403 (1975) (quoting Hall v. Beals, 396 U.S. 45, 49 (1969) (internal quotations omitted))).
79. Id. at 1424. The plaintiffs argued that the lack of safeguards guaranteeing privacy of DNA donors violated their Fourth Amendment right to be free from unreasonable searches and seizures. Id. The plaintiffs also feared that genetic information obtained from the samples would be used to discriminate against applicants for insurance, benefit programs, or jobs. Id.
80. Id. at 1425.
81. Id.
82. Id. at 1426.
83. Id. at 1425.
84. Id. at 1425–26. The changes included shortening the retention period of samples and implementing a mechanism for service members to request destruction of their DNA samples upon separation from the military. Id.
85. 114 F.3d 952 (9th Cir. 1997).
action was “too remote to preserve a live case or controversy.” 86 Valerie Dufresne alleged that California’s use of chemical pesticides against Mediterranean fruit flies exacerbated her Chronic Fatigue Syndrome. 87 She sued several federal and California officials to enjoin the spraying of pesticides to combat future fruit fly infestations. 88 Before the Dufresne case reached the Ninth Circuit, California agriculture officials reported to the court that they had eradicated the fruit fly by using sterile insects, 89 and that they had found no live fruit flies in California since 1995. 90 California officials anticipated that any future outbreak would also be combated with sterile insects. 91 Based on the eradication of the fruit fly and the new, pesticide-free method of combating future infestations, the Dufresne court held that the case was moot. 92

D. Both the U.S. Supreme Court and the Ninth Circuit Allow the Public Interest to Weigh Against Mootness in Determining the Legality of an Action

Federal courts do not recognize a general exception to mootness for cases with a continuing public interest. 93 However, the public interest in having the legality of an issue settled will weigh against mootness provided the case satisfies one of the recognized exceptions to mootness. 94 The U.S. Supreme Court and the Ninth Circuit have considered the public interest in cases applying both the capable-of-

---

86.  Id. at 955.
87.  Id. at 954. A support group for sufferers from Chronic Fatigue Syndrome joined Dufresne’s suit as plaintiffs. Id. at 952.
88.  Id. at 954. Dufresne initially sought monetary damages, but dropped that claim before the case reached Ninth Circuit. Id.
89.  Id.
90.  Id.
91.  Id.
92.  Id. at 955.
An Exception to Mootness

repetition-yet-evading-review and voluntary cessation exceptions to mootness.95

Throughout the U.S. Supreme Court’s development of the exceptions to mootness, the Court has also considered the public interest in settling the legality of a challenged action.96 In Southern Pacific Terminal Co. v. ICC,97 the Court applied the capable-of-repetition-yet-evading-review exception for the first time and considered the public’s interest in settling the legality of a two-year order issued by the Interstate Commerce Commission (ICC).98 Although the order had expired before the case reached the Court, the Court refused to dismiss the case as moot because review of the ICC “ought not to be . . . defeated, by short-term orders, capable of repetition, yet evading review.”99 The Southern Pacific Court went on to quote Boise City Irrigation and Land Co. v. Clark,100 a Ninth Circuit mootness decision, to support its proposition that the case was not moot.101 The Boise City court considered the “propriety of deciding some question of law presented which might serve to guide the municipal body when again called upon to act in the matter.”102

In United States v. W.T. Grant Co.,103 the U.S. Supreme Court held that the voluntary cessation exception,104 “together with a public interest in having the legality of the practices settled, militate[d] against a

95. See W.T. Grant Co., 345 U.S. at 632 (applying the voluntary cessation exception); S. Pac. Terminal Co. v. ICC, 219 U.S. 498, 515–16 (1911) (applying the capable-of-repetition-yet-evading-review exception); Miller ex rel. NLRB v. Cal. Pac. Med. Ctr., 19 F.3d 449, 454 (9th Cir. 1994) (applying the capable-of-repetition-yet-evading-review exception); Armster v. United States Dist. Court for the Cent. Dist. of Cal., 806 F.2d 1347, 1358–59 (9th Cir. 1986) (applying the voluntary cessation exception to mootness).
97. 219 U.S. 498 (1911).
98. Id. at 515–16.
99. Id. at 514–15.
100. 131 F. 415 (9th Cir. 1904). In Boise City, the court held that the appellant’s challenge of a municipal ordinance fixing the water rate did not become moot when the ordinance expired. Id. at 419.
102. Boise City, 131 F. at 419.
103. 345 U.S. 629 (1953).
104. The “voluntary cessation of allegedly illegal conduct does not deprive the tribunal of power to hear and determine the case, i.e., does not make the case moot.” Id. at 632. The voluntary cessation exception prevents a defendant who has ceased the challenged activity from resuming the activity once the case is dismissed as moot. United States v. Concentrated Phosphate Exp. Co., 393 U.S. 199, 202–03 (1968).
mootness conclusion.”105 The federal government brought suit against the director of W.T. Grant Company, who was also the director of three other corporations, for violations of the Clayton Act.106 Because the case was the first opportunity for the Court to consider the Clayton Act’s prohibition of interlocking corporate directors, the Court concluded that there was a public interest in reaching the merits of the case.107

Based on the U.S. Supreme Court’s decisions in *Southern Pacific Terminal Co.* and *W.T. Grant*, the Ninth Circuit has considered the public interest in determining the legality of a challenged action in cases where it has applied both the capable-of-repetition-yet-evading-review and voluntary cessation exceptions.108 After holding that the capable-of-repetition-yet-evading review exception applied in *Miller*, the court stated that “the public interest weighs heavily in favor of our resolving this appeal.”109 The court reasoned that it was important for both the NLRB and employers to know what criteria courts would apply in reviewing preliminary injunctions issued by the NLRB.110

In *Armster v. United States District Court for the Central District of California (Armster II)*,111 the Ninth Circuit considered the public interest in settling the constitutionality of suspending federal civil jury trials.112 The *Armster II* decision followed the court’s earlier decision in *Armster v. United States District Court for the Central District of California (Armster I)*113 that such suspension was unconstitutional.114 In

106. *Id*. at 630.
107. *See id.* at 632.
108. *See Alaska Fish & Wildlife Fed’n & Outdoor Council, Inc. v. Dunkle*, 829 F.2d 933, 939 (9th Cir. 1987) (applying the capable-of-repetition-yet-evading-review exception and citing *W.T. Grant* in support of the proposition that public interest weighs against mootness); *Armster v. United States Dist. Court for the Cent. Dist. of Cal.*, 806 F.2d 1347, 1360 (9th Cir. 1986) (applying the voluntary cessation exception and citing *Southern Pacific Terminal Co.*, *Boise City*, and *W.T. Grant* in support of the proposition that there is a strong public interest when the court decides important precedential issues); *Olagues v. Russoniello*, 797 F.2d 1511, 1517 (9th Cir. 1986) (applying the capable-of-repetition-yet-evading review exception and citing *W.T. Grant* for the proposition that a strong public interest weighs against mootness).
110. *Id*. The Ninth Circuit has also considered a public interest in a wide variety of legal issues. *See, e.g.*, Greenpeace Action v. Franklin, 982 F.2d 1342, 1348 (9th Cir. 1992) (considering the public interest in pollock fishing in the Gulf of Alaska); *Alaska Fish & Wildlife Fed’n*, 829 F.2d at 939 (considering the public interest in Native American hunting of migratory birds); *Olagues*, 797 F.2d at 1517 (considering the public interest in alleged voter rights violations).
111. 806 F.2d 1347 (9th Cir. 1986).
112. *See id.* at 1360–61.
113. 792 F.2d 1423 (9th Cir. 1986).
An Exception to Mootness

June 1986, the Administrative Office of the United States Courts sent a memorandum to all district judges suggesting that they suspend civil jury trials until September 30, 1986 in response to a federal budget shortfall.\(^{115}\) The \textit{Armster I} petitioners sought writs of mandamus from the Ninth Circuit requiring district judges who followed the memo’s suggestion to impanel juries and proceed with civil trials.\(^{116}\) The \textit{Armster I} court held that suspending civil jury trials in response to a budget shortfall violated the Seventh Amendment.\(^{117}\)

On the same day that the \textit{Armster I} court released its decision, Congress approved a supplemental appropriations bill that eliminated the budget shortfall.\(^{118}\) The Administrative Office subsequently rescinded its recommendation to suspend civil jury trials.\(^{119}\) Based on that rescission, the Justice Department filed a motion to vacate the \textit{Armster I} decision for mootness, which was addressed by the Ninth Circuit court in \textit{Armster II}.\(^{120}\) The \textit{Armster II} court rejected the Justice Department’s motion for three reasons.\(^{121}\) First, the court’s holding in \textit{Armster I} addressed the district judges’ suspension of jury trials and not the Administrative Office’s recommendation.\(^{122}\) Consequently, rescission of that recommendation was not a “sufficient basis for mooting [the] decision regarding the constitutional obligation of the . . . district courts.”\(^{123}\) Second, at the time the \textit{Armster I} court rendered its decision, it had before it a case or controversy meeting the Article III justiciability requirements.\(^{124}\) An appellate court is not required to dismiss a case as moot based on events that occur after it has rendered its final judgment.\(^{125}\) Third, the voluntary cessation exception to mootness

\(^{114}\) Id. at 1425.
\(^{115}\) Id. at 1424.
\(^{116}\) Id.
\(^{117}\) Id. at 1425. The \textit{Armster I} court declined to issue the writs of mandamus choosing instead to rely on the district judges to voluntarily reinstate jury trials. Id. at 1431.
\(^{118}\) Armster v. United States Dist. Court for the Cent. Dist. of Cal., 806 F.2d 1347, 1350 (9th Cir. 1986) (finding that Congress approved the Supplemental Appropriations Bill, H.R. 4515, on June 26, 1986); Armster, 792 F.2d at 1423 (submitting decision on June 26, 1986).
\(^{119}\) Armster, 806 F.2d at 1350.
\(^{120}\) Id.
\(^{121}\) Id. at 1353.
\(^{122}\) Id.
\(^{123}\) Id. at 1354.
\(^{124}\) Id.
\(^{125}\) Id. at 1355. Appellate courts do have discretion to vacate earlier decisions, but the \textit{Armster II} court noted no reason to exercise that discretion. Id. at 1355–57.
prevented the case from being moot.\textsuperscript{126} The court held that each of these three grounds was alone sufficient to require it to deny the Justice Department’s motion.\textsuperscript{127}

After applying the voluntary cessation exception, the \textit{Armster II} court reasoned that the strong public interest in determining the constitutionality of suspending jury trials further weighed against mootness.\textsuperscript{128} The court stated that the Ninth Circuit “has long held that there is a strong public interest in the court’s resolving important precedential issues, a public interest that militates against a finding of mootness in cases presenting such issues.”\textsuperscript{129} The court also concluded that “[c]learly, the ‘flexible character of the [Article] III mootness doctrine’ encompasses consideration of the public interest in safeguarding fundamental constitutional rights.”\textsuperscript{130}

In sum, the mootness doctrine recognizes an exception for cases capable-of-repetition-yet-evading-review. The exception is applicable where a challenged action is too short in duration to be fully litigated and when there is a reasonable expectation that the complaining party will be subjected to the same action in the future. Neither U.S. Supreme Court nor Ninth Circuit decisions applying the exception require that there be a demonstrated probability that the challenged action will recur. However, no reasonable expectation exists when the recurrence of the challenged action is “highly unlikely” or “speculative.” Both the U.S. Supreme Court and the Ninth Circuit have bolstered the application of exceptions to mootness with considerations of the public interest in settling the legality of a challenged action.

\section*{II. THE NINTH CIRCUIT HELD \textit{FOSTER V. CARSON} MOOT AND THE CAPABLE-OF-REPETITION-YET-EVADING-REVIEW EXCEPTION INAPPLICABLE}

In response to a severe state budget crisis, the chief justice of the Oregon Supreme Court, acting as the administrative head of the state

\begin{itemize}
\item \textsuperscript{126} \textit{Id.} at 1357.
\item \textsuperscript{127} \textit{Id.} at 1353. When there are two or more grounds on which an appellate court may rest its decision and the court adopts all of those grounds, each of those grounds is the judgment of the court and of equal validity with the other grounds. United States v. Title Ins. & Trust Co., 265 U.S. 472, 486 (1924). None of the grounds are considered dicta. \textit{Id.}
\item \textsuperscript{128} \textit{Armster}, 806 F.2d at 1360.
\item \textsuperscript{129} \textit{Id.}
\item \textsuperscript{130} \textit{Id.} (quoting United States Parole Comm’n v. Geraghty, 445 U.S. 388, 400 (1980)).
\end{itemize}
An Exception to Mootness

judiciary, issued the Budget Reduction Plan (BRP) that suspended appointments of indigent defense counsel between March 1 and June 30, 2003.131 In Foster v. Carson, the plaintiffs brought suit challenging the constitutionality of the suspension.132 Before the Ninth Circuit heard the case, the BRP had expired.133 The Ninth Circuit held that the expiration of the BRP mooted the case and rejected the plaintiffs’ argument that the capable-of-repetition-yet-evading-review exception applied.134

A. In Foster, the Plaintiffs Challenged the Oregon Chief Justice’s Suspension of Indigent Defense Counsel Appointments

The BRP issued by the chief justice cut off funding for the appointment of indigent defense counsel for nine non-violent offenses.135 In response to the BRP, Metropolitan Public Defender Services, Inc., Public Defender Services of Lane County, the District Attorney for Lane County, and several indigent defendants brought suit in state court.136 After the state court dismissed the action, the plaintiffs sued the chief justice in the U.S. District Court for the District of Oregon,137 alleging that the BRP violated their rights under the First, Sixth, and Fourteenth Amendments.138 The plaintiffs requested that the district court declare that the suspension of funding was unconstitutional and void.139 The district court held that all plaintiffs had standing but dismissed the case.

131. Foster v. Carson, 347 F.3d 742, 744 (9th Cir. 2003).
132. Id.
133. Id. at 745.
134. Id. at 745–48.
135. Id. at 744. Offenses included certain misdemeanors, misdemeanor probation violations, adult property and drug felonies, and adult controlled substance possession felonies. Id. The BRP also closed all offices of the Oregon courts on Fridays. Foster, 347 F.3d at 744. At the initial court appearance, affected cases were to be rescheduled for a court appearance in the next budget period, which began on July 1, 2003. Id.
136. Id. at 745.
137. Id. at 742. The defendants also included the state court administrator and several Oregon judges. Id.
138. Id. at 745.
139. Id. at 746. The plaintiffs also sought costs and fees and any other relief that the court deemed appropriate. Id.
on Younger abstention grounds.\textsuperscript{140} Plaintiffs then appealed to the Ninth Circuit.\textsuperscript{141}

B. The Foster Court Held That the Case Was Moot

On appeal, the Ninth Circuit held that the Foster case was moot.\textsuperscript{142} By the time the court heard the case, the chief justice had allowed the BRP to expire, and the Oregon State Legislature had passed a new state budget that removed the short-term need to suspend funds for indigent defense counsel.\textsuperscript{143} With this funding restored, the Foster court concluded that it could neither provide any additional relief to the plaintiffs nor undo the harm caused by the delay in the appointment of counsel.\textsuperscript{144}

C. The Foster Court Also Held That the Capable-of-Repetition-yet-Evading-Review Exception Did Not Apply

The court rejected the plaintiffs’ contention that capable-of-repetition-yet-evading-review exception applied to the case.\textsuperscript{145} The court agreed that because the parties had made all possible efforts to expeditiously litigate the case before the BRP expired, the BRP was too short in duration to be fully litigated before it ceased.\textsuperscript{146} However, the court held that the plaintiffs had not established a reasonable expectation that the chief justice would again cut funding for indigent defense counsel.\textsuperscript{147}

The Foster court held that the Armster II decision provided neither controlling nor persuasive authority for applying the capable-of-repetition-yet-evading-review exception.\textsuperscript{148} In so holding, the court rejected the plaintiffs’ use of the Armster II case as “authority for finding that the . . . exception applies due to the importance of this case.”\textsuperscript{149}

\textsuperscript{140} Id. at 745. Federal courts will not enjoin enforcement of state law unless the facts demonstrate that an injunction is necessary to prevent irreparable harm to the plaintiff. Younger v. Harris, 401 U.S. 37, 43 (1971).
\textsuperscript{141} Foster, 347 F.3d at 745.
\textsuperscript{142} Id. at 746.
\textsuperscript{143} Id. at 745.
\textsuperscript{144} Id. at 746.
\textsuperscript{145} Id. at 748–49.
\textsuperscript{146} Id. at 746.
\textsuperscript{147} Id. at 748.
\textsuperscript{148} Id.
\textsuperscript{149} Id. at 746.
An Exception to Mootness

court distinguished the Armster II decision based on its unique procedural posture.150 After the Armster I case was decided on its merits, the Justice Department moved for dismissal of the case as moot based on the developments that occurred post judgment.151 The Armster II court considered the Justice Department’s motion to dismiss.152

The Foster court explained that Ninth Circuit precedent did not allow the capable-of-repetition-yet-evading-review exception to apply when the plaintiffs demonstrated only “a mere possibility that something might happen.”153 To support this proposition, the court cited Mayfield and Dufresne.154 The court noted that Foster presented a more difficult question of mootness than Dufresne because Oregon’s budget crisis, unlike the Mediterranean fruit fly, had not been eradicated.155 However, the court concluded that the prospect of a future funding suspension, like the possibility of future DNA tests in Mayfield or future pesticide spraying in Dufresne, depended on “speculative contingencies.”156 In Foster, those contingencies included Oregon’s economic condition, indigent defense counsel funding choices made by the legislature, demand for indigent defense counsel, and the chief justice’s reaction to any budget shortfall.157 These unknown contingencies, the court reasoned, outweighed the fact that the chief justice had suspended funding once before.158 The court considered that fact to be the only support for the plaintiffs’ argument that the capable-of-repetition-yet-evading-review exception saved the case from mootness.159 The court placed little weight on the plaintiffs’ contention that the five special legislative sessions cutting funds from the 2001–2003 budget after it had been passed indicated that a passed budget did not guarantee funding.160

150. Id. at 747.
151. Armster v. United States Dist. Court for the Cent. Dist. of Cal., 806 F.2d 1347, 1350 (9th Cir. 1986).
152. Id.; see also supra notes 77–92 and accompanying text.
153. Foster, 347 F.3d at 748.
154. Id.
155. Id.
156. Id. at 748–49.
157. Id. at 748.
158. See id.
159. Id.
160. Id.
D. Information Introduced by the Foster Plaintiffs in a Supplemental Brief Suggested that Oregon’s Budget Crisis Had Not Ended

The Foster plaintiffs submitted a supplemental memorandum to the court suggesting that Oregon’s budgetary crisis was far from resolved and establishing that the Oregon legislature anticipated future funding suspensions for indigent defense counsel.161 The memorandum outlined events that had occurred subsequent to the expiration of the BRP.162 The Oregon legislature balanced the 2003–2005 biennial budget by passing a package of spending and revenue bills that included a controversial three-year income tax surcharge.163 While Foster was before the Ninth Circuit, voters opposing the income tax surcharge were mounting an effort to repeal it through a special referendum.164 In a referendum held in January 2003, Oregon voters rejected a smaller income tax surcharge that the legislature had enacted in response to the severe budget deficit.165 Anticipating that voters might also reject the surcharge used to balance the 2003–2005 biennial budget, the legislature included certain provisions in House Bill 5077166—one of the bills passed to balance the budget—that would disappropriate funds to various state agencies in the event voters repeal the income tax surcharge.167 Among the agencies

161. See Appellants’ Supplemental Memorandum at 1–4, Foster (Nos. 03-035457 & 03-035458).
162. See id. The facts reported by the plaintiffs in their supplemental memorandum are supported by media coverage of the Oregon legislature’s efforts to balance the budget and the resulting voter action to repeal the income tax surcharge. See infra notes 163–168 and accompanying text.
An Exception to Mootness

targeted by House Bill 5077 is the Oregon Judicial Department, which stands to lose $14,414,400 for indigent defense services.\textsuperscript{168}

In sum, the chief justice of the Oregon Supreme Court suspended funds for the appointment of indigent defense counsel. On appeal, the \textit{Foster} court held that the case was moot because the suspension of funds had been lifted. The court also held that the capable-of-repetition-yet-evading-review exception to mootness did not apply because there was no reasonable expectation that the plaintiffs faced injury from a future suspension of funds. In a supplemental brief filed with the court, the \textit{Foster} plaintiffs presented evidence that Oregon’s budget crisis was ongoing and that the Oregon legislature anticipated voter rejection of a key income tax surcharge. Without that surcharge, the Oregon Judicial Department was again at risk of a substantial budget cut.

III. THE CAPABLE-OF-REPETITION-YET-EVADING-REVIEW EXCEPTION PREVENTED \textit{FOSTER} FROM BEING MOOT

Under both U.S. Supreme Court and Ninth Circuit precedent, there was a reasonable expectation in \textit{Foster} that the chief justice would again suspend funding for indigent defense counsel.\textsuperscript{169} While neither court imposes a rigid standard for demonstrating that a reasonable expectation exists, the prevailing Ninth Circuit standard requires that the recurrence of the injury be more than “highly unlikely” or “speculative.”\textsuperscript{170} The facts before the \textit{Foster} court established a reasonable expectation that the action would recur; such recurrence was neither “highly unlikely” nor “speculative.”\textsuperscript{171} The public interest in the legality of the chief justice’s action further supports application of the capable-of-repetition-yet-evading-review exception.\textsuperscript{172}

A. \textit{The Foster Plaintiffs Presented Facts Demonstrating a Reasonable Expectation That the Chief Justice Would Again Suspend Funds for Indigent Defense Counsel}

Neither U.S. Supreme Court nor Ninth Circuit precedent imposes a stringent standard for establishing a reasonable expectation that an

\textsuperscript{168} Or. H.R. 5077, § 88; Appellants’ Supplemental Memorandum at 6.
\textsuperscript{169} See infra Part III.A.
\textsuperscript{170} See supra Part I.B–C.
\textsuperscript{171} See infra Part III.A.
\textsuperscript{172} See infra Part III.B.
action will recur. A party may demonstrate a reasonable expectation by demonstrating “some indication” of recurrence. However, the Ninth Circuit will not apply the capable-of-repetition-yet-evading-review exception when the recurrence of the challenged action is “highly unlikely” or “speculative.” As in both Honig and Miller, the facts in Foster indicated that the challenged action would recur and that such recurrence was neither “highly unlikely” nor “speculative.” The Foster court relied on Mayfield and Dufresne as authority for finding that the capable-of-repetition-yet-evading-review exception did not apply. However, unlike Foster, the facts in both Mayfield and Dufresne demonstrated that the recurrence of the challenged action was “highly unlikely” and “speculative.”

Both the U.S. Supreme Court in Honig and the Ninth Circuit in Miller held that there was a reasonable expectation that a challenged action would recur. In Honig, the U.S. Supreme Court held that there was a reasonable expectation that California would again violate Smith’s rights under the EHA. In order for that to occur, Smith would have to re-enroll in a California school, be placed in an environment where his behavior could not be controlled, act in an aggressive manner caused by his disability, and be expelled for his behavior. The Court concluded that this chain of events constituted a reasonable expectation even though Smith had expressed no intention to re-enroll in a California school and had less than a year remaining before the EHA no longer applied to him. Similarly, in Miller, the Ninth Circuit held that there was a reasonable expectation that CPMC would face a future NLRB

174. Alaska Ctr. for the Env’t v. United States Forest Serv., 189 F.3d 851, 856 (9th Cir. 1999).
175. Unabom Trial Media Coalition v. United States Dist. Court for the E. Dist. of Cal., 183 F.3d 949, 952–53 (9th Cir. 1999); Sze v. INS, 153 F.3d 1005, 1009 (9th Cir. 1998).
176. Dilley v. Gunn, 64 F.3d 1365, 1369 (9th Cir. 1995).
177. See infra notes 189–192 and accompanying text.
178. Foster v. Carson, 347 F.3d 742, 748 (9th Cir. 2003).
179. See infra notes 194–202 and accompanying text.
181. Honig, 484 U.S. at 322–24; see also supra notes 46–61 and accompanying text.
183. See id. at 322–24.
An Exception to Mootness

unfair labor dispute based on its ongoing reorganization, large size, and collective bargaining agreements with five other unions. 184

The facts before the Foster court indicated that the chief justice would again be forced to suspend funding for indigent defense counsel. As part of the 2003–2005 balanced budget, the Oregon legislature approved an income tax surcharge as a necessary source of revenue. 185 The possibility that Oregon voters would reject the surcharge was more than merely speculative—a smaller tax surcharge had been rejected by referendum in January 2003. 186 In response, the Oregon legislature enacted provisions that would automatically inappropriate funds to state agencies in the event that Oregon voters repeal the tax surcharge. 187 These automatic cuts include $14,414,400 from the budget for indigent defense services. 188

In Foster, as in Honig and Miller, the facts provided “some indication” that the challenged action would recur and demonstrated that such recurrence was neither “highly unlikely” nor “speculative.” 189 In Honig and Miller, the recurrence of the challenged action relied on a series of events not guaranteed to occur. 190 In both cases, the courts held that there was a reasonable expectation that the challenged action would recur and applied the capable-of-repetition-yet-evading-review exception. 191 The facts in Foster similarly support a conclusion that the chief justice will again face a funding shortage requiring a suspension of indigent defense counsel appointments. 192

Unlike in Foster, the recurrence of the challenged action in Mayfield was “highly unlikely” and “speculative.” 193 In Mayfield the plaintiffs’

---

184. Miller, 19 F.3d at 454; see also supra notes 69–74 and accompanying text.
185. See supra note 163 and accompanying text.
186. See supra note 165 and accompanying text.
187. See supra note 167 and accompanying text.
188. H.R. 5077, § 88, 72d Legis. Assem., Reg. Sess., 2003 Or. Laws ch. 710, § 88; see also supra note 168 and accompanying text.
189. Compare Honig v. Doe, 484 U.S. 305, 322–24 (1988) (holding that there was a reasonable expectation that the plaintiff would re-enroll in a California school, misbehave because of his disability, and be expelled for that misbehavior) and Miller, 19 F.3d at 454 (holding that it was likely that a hospital would again take action leading to an NLRB proceeding) with supra notes 185–188 and accompanying text (suggesting that Oregon’s ongoing budget crisis will result in a future suspension of indigent defense counsel appointments).
190. See Honig, 484 U.S. at 322–24; Miller, 19 F.3d at 454.
191. See Honig, 484 U.S. at 322–24; Miller, 19 F.3d at 454.
192. See supra notes 163–168 and accompanying text.
suit became moot when both marines left active duty. In addition, because the DOD had changed its DNA sampling program, the marines would not be subjected to the same program even if they were recalled. There was no evidence of a national emergency that would necessitate military recalls nor was there evidence suggesting that the plaintiffs, out of all recently discharged members of the Marine Corps, would be targeted for any such recall. Furthermore, there was no indication that the DOD intended to revert to the version of its DNA sampling program challenged in the original suit. In contrast, the facts presented by the Foster plaintiffs demonstrated that Oregon’s budget crisis would continue and that the state legislature would again cut funding to indigent defense counsel.

The facts in Dufresne also demonstrated that the future use of pesticides was both highly unlikely and speculative. In Dufresne, the effectiveness of the sterile insects—no live Mediterranean fruit flies had been found in California since 1995—and the assertion by California officials that sterile insects would be used to combat any future infestations removed any reasonable expectation that the state would again use chemical pesticides. As the Foster court acknowledged, Oregon’s budget crisis, unlike the Mediterranean fruit fly, had not been “eradicated.” In addition, no Oregon officials asserted that future suspensions of indigent defense counsel would be unnecessary due to an improved method of addressing state budget shortfalls.

The Ninth Circuit has acknowledged that determining whether a case has become moot requires a “highly individualistic . . . appraisal of the
An Exception to Mootness

As in Honig and Miller, the facts before the Foster court established a reasonable expectation that the challenged action would recur. Finding a reasonable expectation in Foster is further supported when the facts in that case are compared with the facts of both the Mayfield and Dufresne cases.

B. The Strong Public Interest in Settling the Legality of Suspending Indigent Defense Counsel Appointments Bolsters Application of the Capable-of-Repetition-yet-Evading-Review Exception to Foster

Under U.S. Supreme Court and Ninth Circuit precedent, courts can weigh the public interest in settling the legality of an issue against mootness. The Ninth Circuit may consider the public interest after finding that either the capable-of-repetition-yet-evading-review or voluntary cessation exceptions to mootness applies. Even though the Foster court correctly held that the Armster II decision did not provide authority for applying the capable-of-repetition-yet-evading-review exception, the Armster II decision provides persuasive authority for allowing the public interest in judging the constitutionality of indigent defense counsel suspensions to weigh against mootness. Based on both U.S. Supreme Court and Ninth Circuit precedent, the Foster court should have allowed the public interest in determining the constitutionality of suspending indigent defense counsel appointments to weigh against mootness.

203. Alaska Ctr. for the Env't v. United States Forest Serv., 189 F.3d 851, 856 (9th Cir. 1999) (quoting 13A CHARLES ALAN WRIGHT & ARTHUR R. MILLER, FEDERAL PRACTICE AND PROCEDURE § 3533 (2d ed. 1984)).
204. See supra notes 163–168, 189–191 and accompanying text; text accompanying note 192.
205. See supra notes 194–202 and accompanying text.
207. See Alaska Fish & Wildlife Fed’n & Outdoor Council, Inc. v. Dunkle, 829 F.2d 933, 939 (9th Cir. 1987) (applying the capable-of-repetition-yet-evading-review exception and citing W.T. Grant in support of the proposition that public interest weighs against mootness); Olagues v. Russoniello, 797 F.2d 1511, 1517 (9th Cir. 1986) (applying the capable-of-repetition-yet-evading review exception and citing W.T. Grant for the proposition that a strong public interest weighs against mootness).
208. Foster v. Carson, 347 F.3d 742, 748 (9th Cir. 2003).
209. See Armster v. United States Dist. Court for the Cent. Dist. of Cal., 806 F.2d 1347, 1360 (9th Cir. 1986).
210. See infra notes 231–239 and accompanying text.
In *Armster II*, the court denied the Justice Department’s motion to vacate the *Armster I* decision as moot on three independently sufficient grounds.211 The court’s second ground for denying the motion was the procedural posture of the case.212 The Justice Department argued that the Administrative Office’s post-judgment rescission of its recommendation to suspend jury trials rendered the case moot.213 The court disagreed and held that it was not required to vacate a judgment as moot based on events that occurred after it had entered that judgment.214 The court went on to hold that the voluntary cessation exception to mootness provided a third ground for denying the Justice Department’s motion.215 After applying the voluntary cessation exception, the court emphasized that the Ninth Circuit “has long held that there is a strong public interest in the court’s resolving important precedential issues, a public interest that militates against a finding of mootness in cases presenting such issues.”216

While the *Foster* court correctly held that the *Armster II* decision was neither controlling nor persuasive authority for applying the capable-of-repetition-yet-evading-review exception, its holding relied on inappropriate grounds.217 The *Foster* court rejected the use of the *Armster II* decision as precedent because the case’s procedural posture differed from that of *Foster*.218 However, the *Armster II* court addressed the case’s procedural posture and held that it was one of the three independent grounds for denying the motion to vacate as moot.219 When an appellate court relies on several independent grounds in its holding, each of those grounds is a valid judgment of the court.220 Therefore, the *Armster II* court’s application of the voluntary cessation exception and the consideration of public interest retain their precedential value even

---

211. *Armster*, 806 F.2d at 1353.
212. *Id.*
213. *Id.* at 1350–51.
214. *Id.* at 1354–57.
215. *Id.* at 1357–61.
216. *Id.* at 1360.
217. *See infra* notes 218–221 and accompanying text.
219. *See Armster*, 806 F.2d at 1350.
220. *United States* v. *Title Ins. & Trust Co.*, 265 U.S. 472, 486 (1924); *see also supra* note 127 and accompanying text.
An Exception to Mootness

though the Armster II case had a different procedural posture than the Foster case.221

The Armster II court’s consideration of the public interest in applying the voluntary cessation exception does not support applying the capable-of-repetition-yet-evading-review exception to Foster based only on the public interest in determining the constitutionality of the chief justice’s actions.222 As the U.S. Supreme Court did in W.T. Grant,223 the Armster II court held that the public interest weighed against a finding of mootness after it had already determined that the voluntary cessation exception applied.224 It did not rely on the public interest alone as grounds for applying the voluntary cessation exception.225 The Foster court was correct when it refused to apply the capable-of-repetition-yet-evading-review exception based only on the public interest.226

Although the Armster II decision does not support using the public interest as the sole basis for applying the capable-of-repetition-yet-evading-review exception to Foster,227 it does support allowing the public interest in determining the constitutionality of a challenged action to weigh against mootness.228 The Armster II court reasoned that it should avoid precluding judicial review of allegedly unconstitutional behavior unless it is “abundantly clear” that it is required to do so by the limitations of the mootness doctrine.229 The court went on to state that “[c]learly, the ‘flexible character of the [Article] III mootness doctrine’ encompasses consideration of the public interest in safeguarding fundamental constitutional rights.”230

The Foster court should have allowed the public interest in determining the constitutionality of the chief justice’s suspension of

221. See Sze v. INS, 153 F.3d 1005, 1009 (9th Cir. 1998) (citing Armster II in applying the voluntary cessation exception); see also Smith v. Univ. of Wash., Law Sch., 233 F.3d 1188, 1194 (9th Cir. 2000) (citing Armster II as authority for the proposition that the voluntariness of cessation is relevant to determining the likely hood of recurrence).
222. See Armster II, 806 F.2d at 1357–61 (considering the public interest after holding that the voluntary cessation exception applied).
224. Armster, 806 F.2d at 1360.
225. Id.
226. See Foster v. Carson, 347 F.3d 742, 748 (9th Cir. 2003).
227. See supra notes 222–226 and accompanying text.
228. Armster, 806 F.2d at 1360–61.
229. Id. at 1360.
230. Id. (quoting United States Parole Comm’n v. Geraghty, 445 U.S. 388, 400 (1980)).
indigent defense counsel appointments to weigh against mootness. Both the U.S. Supreme Court and the Ninth Circuit have recognized that the public interest in determining the legality of a challenged action may militate against mootness. The Ninth Circuit has a long history of considering the public interest in cases involving both the capable-of-repetition-yet-evading-review and voluntary cessation exceptions.

There is a strong public interest in determining the constitutionality of suspending indigent defense counsel appointments similar to the public interest at stake in Armster II. Oregon, along with many of her sister states, continues to face a severe budget crisis that threatens the availability of funds for indigent defense counsel. By holding the case moot, the Foster court avoided deciding whether the Oregon chief justice’s actions were unconstitutional and left him, and officials in other states, free to restrict indigent defendants’ access to counsel in the future. As the Ninth Circuit stated in Boise City, there is “propriety of deciding some question of law presented which might serve to guide the municipal body when again called upon to act in the matter.”

IV. CONCLUSION

The Ninth Circuit should have held that the capable-of-repetition-yet-evading-review exception applied to Foster. Determining whether a case is moot depends on an analysis of the facts in each case, and the facts, as

231. See infra notes 232–239 and accompanying text.
233. See Boise City Irrigation & Land Co. v. Clark, 131 F. 415, 419 (9th Cir. 1904).
234. See Alaska Fish & Wildlife Fed’n & Outdoor Council, Inc. v. Dunkle, 829 F.2d 933, 939 (9th Cir. 1987) (applying the capable-of-repetition-yet-evading-review exception and citing W.T. Grant in support of the proposition that public interest weighs against mootness); Olagues v. RussoSoniello, 797 F.2d 1511, 1517 (9th Cir. 1986) (applying the capable-of-repetition-yet-evading-review exception and citing W.T. Grant for the proposition that a strong public interest weighs against mootness); Armster, 806 F.2d at 1360–61 (applying the voluntary cessation exception and considering the strong public interest in the constitutionality of suspending jury trials).
235. Compare Foster v. Carson, 347 F.3d 742, 744 (9th Cir. 2003) (challenging the constitutionality of suspending indigent defense counsel appointments), with Armster, 806 F.2d at 1350 (challenging the constitutionality of suspending the appointment of juries in federal civil cases).
236. See supra notes 1–2 and accompanying text.
237. See supra notes 163–168 and accompanying text.
238. See Foster, 347 F.3d at 748–49.
239. Boise City Irrigation & Land Co. v. Clark, 131 F. 415, 419 (9th Cir. 1904).
An Exception to Mootness

in *Foster*, may plausibly support different conclusions. However, when examined in light of U.S. Supreme Court and Ninth Circuit mootness precedent, the facts in *Foster* establish a reasonable expectation that the Oregon chief justice will again suspend the appointment of indigent defense counsel. Strong opposition to the income tax surcharge from Oregon voters makes it likely that automatic funding cuts for indigent defense counsel will be triggered.

The U.S. Supreme Court has long recognized that mootness is a more flexible doctrine than standing and that public interest may weigh against dismissing a case as moot. Determining the constitutionality of the chief justice’s actions in *Foster* bolsters the application of the capable-of-repetition-yet-evading-review exception. By refusing to decide the *Foster* case on its merits, the Ninth Circuit sent a powerful and dangerous message to cash-strapped state governments and agencies: short-term suspensions of funding will avoid challenge in federal courts when they expire before the court hears the case. Nothing in the *Foster* decision prevents the Oregon chief justice, or any other judicial official, from suspending funds for indigent defense counsel each time the Oregon Judicial Department faces a budget shortfall.
PRIVATE RAP SHEET OR PUBLIC RECORD?
RECONCILING THE DISCLOSURE OF
NONCONVICTION INFORMATION UNDER
WASHINGTON'S PUBLIC DISCLOSURE AND
CRIMINAL RECORDS PRIVACY ACTS

Lynette Meachum

Abstract: Division I of the Washington State Court of Appeals misapplies Washington’s Criminal Records Privacy Act (CRPA) in determining whether entire files of police investigative information should be available for public review. The plain text and legislative history of the CRPA indicate that the Washington State Legislature intended the CRPA to apply only to the disclosure of criminal history record information, or the data that appears on a subject’s criminal rap sheet. Washington courts should interpret the CRPA narrowly, as an exemption to the broad policy of disclosure established by the state’s Public Disclosure Act (PDA). This approach would reconcile the two statutes in a manner consistent with Washington State Supreme Court precedent. This Comment argues that the Washington State Supreme Court should curtail Division I’s improper application of the CRPA. Specifically, the court should declare that for records subject to public disclosure, the overlap between the CRPA and PDA requires agencies to redact nonconviction criminal history record information while disclosing the remaining record.

Late one night, two police officers respond to a 911 call from an off-duty police officer’s home in Western Washington. The officer’s wife runs outside to meet them, holding her head in obvious pain. Neighbors watching the scene are not surprised at the police’s arrival—they have considered calling 911 when they have overheard previous fights between the couple escalate. The neighbors are, however, surprised when the officers depart after a few minutes and leave the couple alone for the night. One neighbor wonders whether the officers mishandled the events of the evening because a fellow law enforcement officer was involved. The neighbor requests the officers’ investigative reports from the police department under Washington’s Public Disclosure Act (PDA). The chief of police refuses to release the reports. The chief informs her that the department forwarded the case to the county prosecutor, who declined to charge the suspect, and so the matter is

1. Hypothetical created by the author.
2. WASH. REV. CODE §§ 42.17.250–.348 (2000); see id. § 42.17.260(1) (“Each agency, in accordance with published rules, shall make available for public inspection and copying all public records . . . .”)
closed. Under Washington’s Criminal Records Privacy Act (CRPA), the chief argues, nonconviction records may not be released to the public.

If the same situation were to play out in Eastern Washington, the opposite result could occur. Instead of relying on the CRPA to deny the request, the chief there might conclude that the PDA requires the disclosure of investigative reports from a closed case. The chief might point to a state court interpretation of the PDA under which a police report does not enjoy a presumption of privacy after the case is referred to the prosecutor’s office, regardless of the resolution of the case.

The CRPA and PDA reflect Washington’s attempts to satisfy two conflicting principles for the disclosure of publicly held criminal records. When citizens approved the PDA by initiative in 1972, they adopted the philosophy that the public has a right to be fully informed about the actions of its government. The statute’s policy statement reflects this approach:

The people of this state . . . do not give their public servants the right to decide what is good for the people to know and what is not good for them to know. The people insist on remaining informed so that they may maintain control over the instruments that they have created.

According to the Washington State Supreme Court, courts should look to this backbone philosophy of disclosure when interpreting other statutes affecting public records. The PDA’s purpose, however, contrasts with that of the CRPA. Passed five years after the PDA in response to a federal mandate, the CRPA declares that it is state policy to protect the “completeness, accuracy, confidentiality, and security of criminal history record information and victim, witness, and complainant record information.”

4. See id. § 10.97.050; Beltran v. State Dep’t of Soc. & Health Servs., 98 Wash. App. 245, 260, 989 P.2d 604, 612 (1999) (determining that plaintiff should not have been able to obtain investigative information from a case that did not yield a conviction), dismissed upon settlement, 140 Wash. 2d 1021, 10 P.3d 405 (2000).
5. Cowles Publ’g Co. v. Spokane Police Dep’t, 139 Wash. 2d 472, 479, 987 P.2d 620, 623 (1999) (“[T]he fact that allegations have not yet been proven is not persuasive of the need to provide blanket protection for purposes of a defendant’s privacy.”).
6. WASH. REV. CODE § 42.17.251 (emphasis added).
8. WASH. REV. CODE § 10.97.010 (emphasis added).
Disclosure of Nonconviction Information

As the above hypothetical illustrates, the PDA and CRPA conflict over the disclosure of information about law enforcement investigations that have failed to yield a conviction. In interpreting the two statutes, Division I of the Washington State Court of Appeals has applied the CRPA to entire bodies of investigative information—potentially allowing the use of the CRPA to exempt all investigative information from disclosure when it concerns a nonconviction. Investigative information can include all records prepared by the police in connection with an investigation, such as arrest reports, patrol reports, officer reports, and citation information. Although the Washington State Supreme Court has not ruled on Division I’s approach, its cases can support an alternative construction of the CRPA. Under this alternative construction, the CRPA does not remove entire investigative files from the reach of the PDA; instead, it requires the redaction of a narrow category of protected information before a law enforcement agency (such as a police department) may disclose those records.

This Comment argues that the Washington State Supreme Court should correct Division I’s erroneous application of the CRPA to entire bodies of investigative information. The PDA and its limited privacy rights more appropriately govern the disclosure of investigative records. The CRPA should govern only “criminal history record information”—a narrow category of compiled information typically found on the subject’s rap sheet. Criminal history record information consists only of a name, a notation of that person’s contact with law enforcement, and the disposition of the incident. This approach is consistent with the text of the CRPA. Further, neither federal nor state legislative history suggests that legislators intended the CRPA’s reach to extend to investigative information.

14. See infra Part V.
16. See id. § 10.97.030(1).
17. Id.
18. See id. § 10.97.050.
19. See infra Part II.B.
cases recognize the need to exclude information protected by the CRPA from investigative information that is otherwise subject to disclosure under the PDA. However, the higher court has not laid to rest Division I’s expansive interpretation of the scope of that protection.

Part I of this Comment explores the PDA’s governance of public records. Part II addresses the privacy right to criminal history record information created by the CRPA. Part III provides an overview of the Washington State Supreme Court’s approach to the disclosure of investigative information, and Part IV discusses the conflicting approach of Division I of the Washington State Court of Appeals. Part V then argues that the Washington State Supreme Court should curtail Division I’s inappropriate application of the CRPA to investigative information.

I. THE PDA GOVERNS ACCESS TO PUBLIC RECORDS IN WASHINGTON STATE

Washington’s broad PDA allows public access to all records that relate to the conduct of government, subject to a list of narrow exemptions. The PDA exempts police investigative reports only when their disclosure would harm effective law enforcement or any person’s right to privacy. Although the Washington State Legislature can further define privacy rights that exist within the PDA, the Washington State Supreme Court has ruled that the statute’s underlying policy of disclosure requires courts to interpret any exemptions to the PDA narrowly.

22. WASH. REV. CODE § 42.17.260(1).
23. Id. § 42.17.310.
24. Id. § 42.17.310(1)(d).
Disclosure of Nonconviction Information

A. The PDA Makes All Public Records Available to the Public, Subject to Certain Narrowly Tailored Exemptions

With narrow exceptions, the PDA declares that any record related to the conduct of government and used by a government agency is available to the public. The statute defines public records as “any writing containing information relating to the conduct of government or the performance of any governmental or proprietary function prepared, owned, used, or retained by any state or local agency regardless of physical form or characteristics.” Agencies cannot use the presence of private information in a record as a justification for refusing to release the entire record. Rather, the PDA requires that officials redact private information and release the remainder of the record.

One of the PDA’s limited exemptions addresses law enforcement investigative records. Officials may refuse to disclose investigative records if withholding the records is essential to (1) effective law enforcement, or (2) the protection of any person’s right to privacy. The PDA instructs courts to construe this right to privacy narrowly, as with every exemption to the PDA. In fact, the Washington State Legislature amended the PDA to emphasize that courts are to focus on the statute’s broad mandate for disclosure of public records. The PDA’s construction clause highlights the narrow character of its exemptions.

26. Section 42.17.310 of the Revised Code of Washington includes fifty-six separate exemptions. WASH. REV. CODE § 42.17.310(1)(a)–(ggg). Relevant to this Comment are sections 42.17.310(1)(d) (exempting certain investigative records), 42.17.310(1)(c) (exempting taxpayer information when disclosure would violate the taxpayer’s right to privacy), and 42.17.310(1)(j) (exempting work product when sought as part of pretrial discovery). See id. § 42.17.310(1); see also Dawson v. Daly, 120 Wash. 2d 782, 789–90, 845 P.2d 995, 1000 (1993).

27. WASH. REV. CODE § 42.17.020(36) (defining the term “public records”).

28. Id. § 42.17.260(1).

29. Id. § 42.17.020(36).

30. Id. § 42.17.260(1).

31. Id.

32. Id. § 42.17.310(1)(d).

33. Id.

34. Id. § 42.17.251; see also Amren v. City of Kalama, 131 Wash. 2d 25, 31, 929 P.2d 389, 392 (1997); Progressive Animal Welfare Soc’y v. Univ. of Wash., 125 Wash. 2d 243, 251, 884 P.2d 592, 597 (1994).


36. Id.
instructing courts to take into account the statute’s policy of free and open examination of public records.\textsuperscript{37}

B. Courts Interpret Subsequent Laws Altering Privacy Rights Within the PDA as Subject to the PDA’s Broad Policy Favoring Disclosure

The Washington State Supreme Court has established that courts should attempt to reconcile subsequent laws that alter the PDA’s disclosure provisions with the PDA’s preference for disclosure.\textsuperscript{38} In \textit{Hearst Corp. v. Hoppe},\textsuperscript{39} the court considered whether a 1973 law that specifically exempted certain taxpayer income information from public inspection\textsuperscript{40} amended the PDA.\textsuperscript{41} The PDA already protected the release of such data under the provision that exempted disclosure violating a taxpayer’s right to privacy.\textsuperscript{42} Relying on a rule to treat statutes dealing with the same subject as parts of “a harmonious total schema,”\textsuperscript{43} the court held that the later law should be construed as a more specific, supplemental rule to the taxpayer privacy right existing within the PDA.\textsuperscript{44} The court did not find the later law to be a “significant departure” from the PDA, although each involved a different disclosure standard.\textsuperscript{45} Reconciling the two statutes into a single scheme, the court concluded that courts must read the PDA’s “strongly-worded mandate for broad disclosure of public records” into the later law.\textsuperscript{46} Relying on the PDA’s policy statement, the explanation of the statute from the voter’s pamphlet, and analogous disclosure policies at the federal level,
Disclosure of Nonconviction Information

the court noted that all of these sources pointed to the law’s “expansive disclosure requirement” and should serve as a guide for reconciling later laws that affect disclosure.47

The *Hearst* court also interpreted the scope of a privacy right within the PDA for the first time, giving it narrow limits48 that have since become the standard for interpreting all of the statute’s privacy rights.49 Because the public records section of the PDA did not initially define the term “privacy,”50 the *Hearst* court turned to the common-law test for invasion of privacy.51 The court adopted the *Restatement (Second) of Torts*’ test that information is private only if its disclosure (1) would be highly offensive to a reasonable person, and (2) is not of legitimate concern to the public.52 Washington lawmakers ultimately codified the *Hearst* court’s application of the common-law invasion of privacy test.53

In sum, Washington’s PDA requires disclosure of all public records not subject to an exemption.54 Agencies and courts must narrowly interpret all exemptions,55 including the limited exemption for certain police investigative reports.56 Further, the PDA’s privacy exemptions extend only to situations where disclosure of the records would be highly offensive to a reasonable person and the information is not of legitimate concern to the public.57 Although legislators may make additional information private, courts are to reconcile new exemptions to the PDA with the statute’s broad policy favoring disclosure.58

47. See id. at 128, 438 P.2d at 249.
48. Id. at 135–36, 580 P.2d at 253.
52. *Hearst*, 90 Wash. 2d at 135–36, 580 P.2d at 253; *RESTATEMENT (SECOND) OF TORTS* § 652D.
53. WASH. REV. CODE § 42.17.255.
54. Id. § 42.17.260(1).
55. Id. § 42.17.251.
56. Id. § 42.17.310(1)(d).
57. Id. § 42.17.255; see also *Hearst*, 90 Wash. 2d at 135–36, 580 P.2d at 253.
II. THE CRPA GOVERNS THE RELEASE OF CRIMINAL HISTORY RECORD INFORMATION

The CRPA restricts the circumstances in which criminal-justice agencies may disseminate criminal history record information. The CRPA defines criminal history record information to include the combination of data typically contained in a person’s rap sheet: a name, contacts with law enforcement, and the disposition of such incidents. The statute places the most restrictions on dissemination of information not involving a conviction. Passed in response to a federal mandate, the CRPA closely resembles the model federal law.

A. The CRPA Prevents Disclosure of Nonconviction Information Contained in a Subject’s Criminal History

The CRPA contains regulations for disseminating criminal history record information. The statute defines criminal history record information as the combination of three necessary elements: (1) identifiable descriptions of a person, (2) notations of arrests or formal criminal charges, and (3) the dispositions of such charges. Law enforcement agencies across the country compile this information about individual suspects and exchange the compilations with other agencies in response to criminal background checks. Agencies typically refer to these compilations as “rap sheets.” The same criminal history record information may also appear in other documents, such as investigative reports. The CRPA specifically excludes wanted posters, records of minor traffic violations, and other police records maintained chronologically rather than by name (such as the daily jail log) from the definition of criminal history record information.

59. WASH. REV. CODE § 10.97.050.
60. Id. § 10.97.030(1).
61. Id. § 10.97.050(3)–(5).
63. WASH. REV. CODE § 10.97.050.
64. Id. § 10.97.030(1).
65. 28 C.F.R. §§ 20.3(l)–(p), 20.33.
66. Id. § 20.3(d).
67. Id. pt. 20 app. (explaining the definition of criminal history record information contained in § 20.3(d)).
68. WASH. REV. CODE § 10.97.030(1).
Disclosure of Nonconviction Information

Whether the public may access criminal history record information depends on its classification. The information can take one of three forms: conviction records, information on current investigations, and nonconviction data. Conviction records relate to incidents with a disposition adverse to the suspect, such as a conviction, a guilty plea, or a dismissal after a period of probation or deferral of sentence. The CRPA gives agencies permission to disseminate conviction records without restriction. The CRPA also allows law enforcement agencies to disseminate information on current investigations as they see fit. The final category, nonconviction data, includes criminal history record information that has not led to a disposition adverse to the suspect and for which proceedings are no longer actively pending.

The CRPA places the most restrictions on the disclosure of nonconviction data. If a case is closed without a conviction, an agency may disclose the criminal history record information only if disclosure is (1) specified by other statutes; (2) necessary for the provision of criminal justice services; or (3) related to research, evaluative, or statistical purposes. The CRPA does not specifically prohibit the dissemination of nonconviction data in other circumstances. However, its provisions allowing dissemination of conviction records and information from current investigations “without restriction” precede the list of circumstances in which nonconviction information may be disclosed. The list therefore defines the parameters of disclosure of nonconviction data.

69. Id. § 10.97.030(3).
70. Id. § 10.97.050(2).
71. Id. § 10.97.030(2).
72. Id. § 10.97.030(3).
73. Id. § 10.97.030(4).
74. Id. § 10.97.050(1).
75. Id. § 10.97.050(2).
76. Id. § 10.97.030(2). There is a rebuttable presumption that proceedings are no longer actively pending when a case remains open for a year without a disposition. Id.
77. Id. § 10.97.050.
78. Id. § 10.97.050(4).
79. Id. § 10.97.050(3), (5).
80. Id. § 10.97.050(6).
81. Id. § 10.97.050(1)-(2).
Individuals have the right to view their own criminal history record information, regardless of whether that information concerns current investigations, convictions, or nonconviction incidents. However, individuals may reproduce their own records only to challenge the accuracy of the information. Furthermore, the inspection right does not include access to investigative files or any information other than criminal history record information. In other words, although an agency must allow individuals to view their nonconviction criminal history record information, this right of access does not extend beyond the three components of the CRPA’s definition of criminal history record information.

B. The CRPA Closely Parallels the Federal Regulations That Mandated Its Passage

Washington legislators adopted the CRPA in response to a federal mandate. Reacting to privacy concerns, particularly in light of the increasing use of computers to share criminal justice information, the U.S. Department of Justice in 1975 issued rules requiring states that received certain federal funding to adopt plans to ensure the accuracy of criminal history record information and to limit its dissemination. The federal government supplied a set of model regulations for this purpose and threatened states with fines for noncompliance, as well as the loss of federal funds. The funding at issue supported the record-keeping systems that local law enforcement agencies used to share information.


83. WASH. REV. CODE § 10.97.080.
84. Id.
85. Id.
86. Id.
87. STAFF OF SENATE JUDICIARY COMM., ANALYSIS OF SB 2608, 45th Leg., 1st Ex. Sess., at 1 (Wash. 1977) (analyzing the bill as of March 22, 1977).
88. TOM DALTON, ANALYSIS OF SB 2608, 45th Leg., 1st Ex. Sess., at 1 (Wash. 1977) (analyzing the bill as of April 11, 1977).
90. Id. § 20.21.
91. Id. pt. 20.
92. Id. § 20.25.
93. Id.
Disclosure of Nonconviction Information

In response to this mandate, Washington passed the CRPA.\(^\text{94}\) During the legislative process, the Senate Judiciary Committee expressly eliminated provisions that exceeded the mandate of the federal regulations.\(^\text{95}\) Washington’s CRPA parallels the model federal regulations more closely than analogous laws passed in other states.\(^\text{96}\) In particular, state legislators adopted the definition of criminal history record information found within the federal regulations.\(^\text{97}\) The CRPA also limited privacy protections to those in the federal model.\(^\text{98}\)

C. The CRPA Affects Information Already Governed by the PDA

The text of the CRPA indicates that it applies to information that may already be subject to disclosure under the PDA.\(^\text{99}\) Although the PDA permits copying of public records for any purpose,\(^\text{100}\) the CRPA specifically states that individuals inspecting their own criminal history records cannot rely on the PDA to copy those records for any purpose other than to challenge the records’ accuracy.\(^\text{101}\) The CRPA also expressly trumps the PDA in its construction clause—nothing in the PDA’s list of exemptions may preclude dissemination of criminal

---

94. STAFF OF SENATE JUDICIARY COMM., supra note 87, at 1 (“Senate Bill 2608 has been drafted by the advisory committee to implement the state plan in order to bring the state into conformance with the requirements of the federal regulations.”).

95. STAFF OF SENATE JUDICIARY COMM., ANALYSIS OF SB 2608, 45th Leg., 1st Ex. Sess., at 1 (Wash. 1977) (analyzing the bill as of May 3, 1977).


97. Compare WASH. REV. CODE § 10.97.030(1) (“Criminal history record information’ means information contained in records collected by criminal justice agencies, other than courts, on individuals, consisting of identifiable descriptions and notations of arrests, detentions, indictments, informations, or other formal criminal charges, and any disposition arising therefrom . . . .”) with 28 C.F.R. § 20.3(d) (“Criminal history record information means information collected by criminal justice agencies on individuals consisting of identifiable descriptions and notations of arrests, detentions, indictments, informations, or other formal criminal charges, and any disposition arising therefrom . . . .”).


99. See WASH. REV. CODE §§ 10.97.080, .140.

100. See id. §§ 42.17.260(1), .270.

101. Id. § 10.97.080.

102. Id. § 10.97.140.

103. Id. § 42.17.310.
D. Several Interpretations of the CRPA Distinguish Investigative Information from the Criminal History Record Information Protected by the CRPA

While the text of the CRPA does not draw a specific distinction between the criminal history record information that constitutes a rap sheet and the broader information compiled by law enforcement officials in the course of an investigation, its federal predecessor and other authorities have distinguished between the two. Information from Washington’s Municipal Research & Services Center indicates that although some interpretations apply the CRPA to entire police reports, the “common interpretation of the [CRPA] is that it applies to criminal histories.”

The Washington State Office of the Attorney General has also indicated that the release of police investigative information is governed only by the PDA, noting in its Open Records & Open Meetings deskbook that “[i]nvestigative information does not fall within the definition of ‘criminal history record information’” as addressed by the CRPA. The comments to the federal regulations on which the CRPA is based also specify that the regulations do not extend to investigative information in police reports.

The U.S. Supreme Court has distinguished a rap sheet’s compilation of criminal history from other investigative information when

104. Id. § 10.97.140. Conviction records, for example, are subject to disclosure without restriction under the CRPA. Id. § 10.97.030(1).

105. The CRPA’s construction clause was added in 1999, within a bill that also required permanent retention of investigative records relating to sex offenders. See Sexually Violent Offenses—Records, ch. 326, sec. 4, 1999 Wash. Laws 1696. The bill created a new exemption to the PDA for these records, and also amended the CRPA to indicate that criminal history record information appearing in such records is still subject to disclosure when the CRPA requires it. Id.


Disclosure of Nonconviction Information

interpreting federal public disclosure requirements. 109 In United States Department of Justice v. Reporters Committee for Freedom of the Press, 110 the Court noted that rap sheets contain accumulated private information about individuals but reveal little or nothing about the conduct of a government agency. 111 They therefore differ from reports that detail agency conduct, 112 such as police reports that show actions taken by law enforcement officers in the course of an investigation. The Court recognized the difference between the scattered disclosure of information within investigative reports and court records, and the disclosure of a compiled rap sheet. 113 The Court held that rap sheets were private information outside the purview of the federal Freedom of Information Act. 114 The Washington State Supreme Court has held that such judicial interpretations of the federal Freedom of Information Act are helpful in construing the State’s PDA. 115

In sum, the CRPA regulates the disclosure of criminal history record information, regardless of whether that information is subject to disclosure under the PDA. 116 The text of the CRPA and the federal regulations that mandated it indicate that the CRPA applies only to the combination of information that would appear on a rap sheet: name, arrest information, and the disposition of that arrest. 117 Washington state officials and commentators 118 and the U.S. Supreme Court 119 have also drawn a distinction between police reports, which are investigative information, and the brief criminal histories compiled on rap sheets.

111. Id. at 773–74 (holding that disclosure of an FBI rap sheet constituted an unwarranted invasion of privacy because it did not involve information about a government agency, but records about a private citizen).
112. Id.
113. Id. at 764.
114. Id.
115. Hearst Corp. v. Hoppe, 90 Wash. 2d 123, 128, 580 P.2d 246, 249 (1978) (“The state act closely parallels the federal Freedom of Information Act . . . , and thus judicial interpretations of that act are particularly helpful in construing our own.”).
116. See supra Part II.C.
118. See WASH. STATE OFFICE OF THE ATTORNEY GEN., supra note 107, § 5.3; Horst, supra note 106, at 5-3 n.1.
Although individuals may view their criminal history record information on rap sheets, the CRPA prohibits the release of that information to the general public when it concerns a contact with law enforcement that fails to yield a conviction.

III. BOTH THE PDA AND CRPA AFFECT THE RELEASE OF POLICE INVESTIGATIVE INFORMATION

The Washington State Supreme Court applies the PDA broadly to police investigative information. The court has indicated that the PDA allows the release of investigative information once the investigation is complete. The court has also affirmed the application of the CRPA to differentiate protected information from investigative information which is otherwise subject to disclosure under the PDA.

A. The Washington State Supreme Court Has Held That Investigative Information Does Not Enjoy a Presumption of Privacy Once It Has Been Submitted for Charging

Under the PDA, investigative records do not enjoy a presumption of privacy once the investigation is complete, regardless of its ultimate disposition. Although the Washington State Supreme Court initially extended a categorical exemption from disclosure to the records of all active cases, the court has since held that this categorical exemption ends when a case is referred for prosecution. In doing so, the court

120. See supra notes 83–86 and accompanying text.
121. See supra notes 77–82 and accompanying text.
123. Id. at 479, 987 P.2d at 623.
125. Newman v. King County, 133 Wash. 2d 565, 574, 947 P.2d 712, 716 (1997) (“We hold the broad language of the statutory exemption requires the nondisclosure of information compiled by law enforcement and contained in an open and active police investigation file because it is essential for effective law enforcement.”).
126. Cowles, 139 Wash. 2d at 479–80, 987 P.2d at 623 (“[W]e hold in cases where the suspect has been arrested and the matter referred to the prosecutor, any potential danger to effective law enforcement is not such as to warrant categorical nondisclosure of all records in the police investigative file.”).
Disclosure of Nonconviction Information

indicated that an unresolved investigation is not itself persuasive of a right to privacy in the investigative information. 127

For a short period in the 1990s, a Washington State Supreme Court decision severely limited public access to investigative records. In *Newman v. King County*, 128 the court held that the PDA categorically exempted from disclosure all investigative information pertaining to open cases. 129 The *Newman* case involved a reporter’s request for access to the King County Department of Public Safety file on the twenty-five-year-old, unsolved murder of civil rights leader Edwin Pratt. 130 The county declined to release anything beyond the initial incident report, citing the PDA exemption for records essential to effective law enforcement. 131 The county also argued that disclosure would violate the privacy rights of witnesses, suspects, and defendants. 132

The Washington State Supreme Court agreed with the county and established a categorical exemption from the PDA for disclosure of “open and active” police files. 133 The court based its holding on the PDA’s effective law enforcement exemption, however, and did not address the county’s privacy argument under the PDA. 134 In addition, the court did not address whether the CRPA would affect the privacy of the investigative information.

Two years later, the Washington State Supreme Court held that the *Newman* decision’s categorical exemption for effective law enforcement ends when a case is referred to the prosecutor for a charging decision. 135 The newspaper plaintiff in *Cowles Publishing Co. v. Spokane Police Department* 136 requested the incident report and booking photo from the DWI arrest of an assistant city attorney. 137 On the same day as the arrest,

127. *Id.* at 479, 987 P.2d at 623 (“[T]o the extent protection of the trial process or the privacy rights of a suspect are essential in any given case, the trial court should make that factual determination on a case-by-case basis.”).
129. *Id.* at 574, 947 P.2d at 716.
130. *Id.* at 568, 947 P.2d at 713.
131. *Id.* at 570, 947 P.2d at 714 (citing WASH. REV. CODE § 42.17.310(1)(d) (2000)).
132. *Id.* at 572, 947 P.2d at 715.
133. *Id.* at 574, 947 P.2d at 716.
137. *Id.* at 474–75, 987 P.2d at 621.
the police department referred the case to the prosecutor and requested that charges be filed.\textsuperscript{138} The court held that once investigators submit a case for a charging decision, withholding disclosure of the investigation is no longer essential to effective law enforcement.\textsuperscript{139} The court noted that “[i]n such circumstances, the risk of inadvertently disclosing sensitive information that might impede apprehension of the perpetrator no longer exists.”\textsuperscript{140}

Unlike the \textit{Newman} decision, the court in \textit{Cowles} also directly addressed the suspect’s privacy interest in criminal investigative information, though only with respect to the PDA.\textsuperscript{141} The court held that nothing in the police report implicated the privacy prong of the PDA’s investigative records exemption.\textsuperscript{142} Emphasizing that the public is “well aware” of the presumption of innocence, the court stated that the fact that allegations are unproven does not alone justify a blanket protection for privacy.\textsuperscript{143}

The \textit{Cowles} decision did not discuss the CRPA. Although the CRPA may implicate the privacy of criminal history record information contained within investigative records, the trial court had held that the CRPA did not prevent the release of the records,\textsuperscript{144} and the city did not appeal that issue.\textsuperscript{145} However, the court’s interpretation of the PDA shows the limited nature of a person’s right to privacy in any investigative information.

\begin{footnotesize}
\begin{footnotes}
\textsuperscript{138.} Id. at 475, 987 P.2d at 621.
\textsuperscript{139.} Id. at 479–80, 987 P.2d at 623.
\textsuperscript{140.} Id. at 477–78, 987 P.2d at 622.
\textsuperscript{141.} Id. at 479, 987 P.2d at 623.
\textsuperscript{142.} Id. at 480, 987 P.2d at 624.
\textsuperscript{143.} Id.
\textsuperscript{144.} Cowles Publ’g Co. v. Spokane Police Dep’t, No. 97-2-03783-7 (Super. Ct. Spokane County Aug. 25, 1997) (judgment and order requiring disclosure of public records).
\textsuperscript{145.} \textit{See} Appellant’s Brief, Cowles Publ’g Co. v. Spokane Police Dep’t, 92 Wash. App. 1018 (1998) (No. 16870-1-III). Because the CRPA permits law enforcement agencies to disclose information about current investigations at their discretion, the privacy right created by the CRPA would not affect records of an open investigation, such as in \textit{Cowles}. \textit{See} WASH. REV. CODE § 10.97.050(2) (2000).
\end{footnotes}
\end{footnotesize}
Disclosure of Nonconviction Information

B. The Washington State Supreme Court Has Peripherally Addressed the CRPA’s Construction, Leaving the Extent of Its Right to Privacy Unclear

The Washington State Supreme Court has indicated that the CRPA affects the release of investigative information under the PDA, but has not clarified the limits of the CRPA’s right to privacy. In *Barfield v. City of Seattle*, the court held that police internal investigation files were not privileged from disclosure under the PDA. The defendants sought the files during discovery in two civil suits against police officers. In both cases, the court determined that the police department had not shown that disclosure of the files would violate anyone’s right to privacy under the PDA’s investigative records exemption. In addition, the court cited the CRPA in affirming the trial court’s decision to delete from the files “all criminal records.”

In short, the court interpreted the CRPA to provide a right to privacy that exempted information otherwise subject to disclosure under the PDA. However, the court did not clarify whether it considered “all criminal records” to include any information related to a criminal case, or only criminal history record information as defined by the CRPA. The decision did recognize that the PDA and CRPA impose different requirements on investigative information. Although the *Barfield* court failed to explain precisely what type of information agencies should withhold, the case establishes that agencies can reconcile the two statutes by separating information that fits under the CRPA’s narrower privacy right from information that does not.

The court has also affirmed a similar interpretation of the CRPA by Division II of the Washington State Court of Appeals, but again the high court left the extent of the CRPA’s protection unclear. In *Limstrom v.*

---

147. *Id.* at 885, 676 P.2d at 442 (noting that the files were, however, subject to a protective order by the court).
148. *Id.*
149. *Id.*
150. *Id.*
151. *See id.*
152. *Id.*
153. *See id.*
Ladenburg, a defense attorney sought prosecution files that contained investigative records involving a certain sheriff’s deputy. Division II reversed a trial court’s decision that the PDA’s work-product exemption protected the records from disclosure. The appellate court also noted that it was improper for the prosecutor’s office to make a blanket claim of exemption for the records under the CRPA because the office made no attempt to withhold only nonconviction information. On appeal, the Washington State Supreme Court agreed. The court remanded the case with instructions for the trial court to determine whether the CRPA protected some of the information from disclosure. While indicating that the CRPA protected the nonconviction information among the requested records, the court again failed to specify whether that protected information included investigative records, or merely rap sheet information.

The Washington State Supreme Court’s holdings in both Barfield and Limstrom support the conclusion that data protected by the CRPA should not be disclosed along with other public records. However, the court has left the scope of the information protected by the CRPA unclear. In a footnote to its Limstrom opinion, the court recognized that it has not yet considered the interpretation of the CRPA adopted by Division I of the Washington State Court of Appeals. Stating that the Limstrom case presented an insufficient record, the Washington State Supreme Court refused to address the issue.

IV. DIVISION I HAS RELIED ON THE CRPA TO GOVERN DISCLOSURE OF ENTIRE INVESTIGATIVE FILES

Division I of the Washington State Court of Appeals has applied the CRPA not only to criminal history record information as defined by the

156. Id. at 528, 933 P.2d at 1057.
157. Id. at 532, 933 P.2d at 1059.
158. Id. at 532–33, 933 P.2d at 1059.
160. Id. at 615–16, 963 P.2d at 879–80.
161. Id. at 616, 963 P.2d at 880.
162. Id. at 616 n.10, 963 P.2d at 880 n.10.
163. Id.
Disclosure of Nonconviction Information

statute, but also to entire bodies of investigative information. This approach differs from that set forth by the Washington State Supreme Court in the Barfield and Limstrom decisions. In those cases, the court affirmed the exclusion of information protected by the CRPA from otherwise public investigative information. The higher court, however, has not reviewed the cases in which Division I interpreted the CRPA.

A. Division I Applied the CRPA to an Entire Investigative File in Hudgens v. City of Renton

In Hudgens v. City of Renton, Division I of the Washington State Court of Appeals determined that the CRPA’s provisions applied to an entire investigative file. “Free-lance newsman” Harley Hudgens requested access to every record prepared by the Renton Police Department in connection with the DWI arrest of a suspect who had been tried and found not guilty. Because of the acquittal, criminal history record information about the case fit the CRPA’s definition of nonconviction information. Division I’s holding, however, extended the CRPA’s protection to the entire investigative file, not just the three elements of criminal history record information.

The Hudgens court, however, limited the reach of the CRPA in one important respect. The court held that the CRPA prevented the copying of the documents, but did not prohibit their inspection by a third party.

166. See Barfield v. City of Seattle, 100 Wash. 2d 878, 885, 676 P.2d 438, 442 (1984); Limstrom, 136 Wash. 2d at 616, 963 P.2d at 880.
169. Id. at 843, 746 P.2d at 321–22.
170. Id. at 843, 746 P.2d at 321.
171. Id.
174. See WASH. REV. CODE § 10.97.030(1).

The court’s analysis relied on a close reading of the CRPA’s “Inspection of Information by Subject” provision, which specifies that “[n]o person shall be allowed to retain or mechanically reproduce any nonconviction data.” Based on this provision, the court concluded that “nothing in the statute prohibits the viewing or inspection of nonconviction information.”

Although the investigative file received protection from copying under the CRPA, the court held that the PDA’s investigative records exemption did not provide any supplemental privacy. The court stated that because the city had not met its burden of showing that release of the file would be highly offensive and outweigh the public’s interest in disclosure, the information was not exempt from disclosure. The court therefore held that all of the information in the investigative file should be accessible for viewing under the PDA, but could not be reproduced under the CRPA.

Thus, the Hudgens decision expanded the scope of information covered by the CRPA while simultaneously decreasing the level of protection it provides. The decision broadened the application of the CRPA to entire investigative files, but limited its protection to preventing only physical retention or copying of protected information. By applying the PDA’s privacy standard, the court preserved the PDA’s broad mandate for access to public records. In the wake of Hudgens, a Seattle assistant city attorney opined that cities can “skip analysis under CRPA and analyze a request for non-conviction data under the provisions of the PDA.” The Washington State Supreme Court denied review of the Hudgens decision, but has

176. Id. at 844, 746 P.2d at 321.
177. WASH. REV. CODE § 10.97.080.
179. Id. at 846, 746 P.2d at 322.
180. Id.
181. Id. at 844–46, 746 P.2d at 321–22.
182. Id. at 843–44, 746 P.2d at 321.
183. Id. at 844–45, 746 P.2d at 321.
184. See id. at 845–46, 746 P.2d at 322; see also WASH. REV. CODE § 42.17.010(11) (2000).
Disclosure of Nonconviction Information

indicated that it may reconsider the appellate court’s interpretation of the CRPA if provided with a sufficient record.187

B. Applying the CRPA in Beltran v. State Department of Social and Health Services, Division I Left It Unclear Whether Access to Nonconviction Investigative Information Is Ever Permissible

Division I of the Washington State Court of Appeals has since followed the Hudgens decision by applying the CRPA to investigative information.188 However, in doing so, the court has sealed this kind of information instead of making it accessible.189 In Beltran v. State Department of Social and Health Services,190 the plaintiff sought to use investigative information about nonconviction criminal activity in a foster home to show that the state was negligent in licensing the foster parent.191 The court held that the dissemination of such information by the Department of Social and Health Services to a third party would violate the foster parent’s privacy right under the CRPA.192 The court affirmed the trial court’s order to seal the investigative information and strike it from the trial record.193

The court’s decision in Beltran left it uncertain whether the CRPA permits the public to access nonconviction investigative information at all.194 Noting that the Hudgens decision prohibited the reproduction of investigative information, the court held that the Beltran plaintiff’s reproduction for evidentiary purposes was improper.195 However, the court’s language did not clarify whether the plaintiff, a third party not involved in the police investigation, should even have been allowed to view the investigative information.196 "Beltran does not fall within any

187. Limstrom v. Ladenburg, 136 Wash. 2d 595, 616 n.10, 963 P.2d 869, 880 n.10 (1998) ("Without a sufficient record, we decline to consider whether we agree with the interpretation of [the CRPA] as articulated in Hudgens . . . .").
189. Id.
191. Id. at 259, 989 P.2d at 611.
192. Id. at 259–60, 989 P.2d at 612.
193. Id. at 260, 989 P.2d at 612.
194. See id. at 259, 989 P.2d at 612.
195. Id. at 260, 989 P.2d at 612.
196. See id.
of the exceptions provided in the Act for those who are allowed access to this type of information,” the court wrote.197 The court held that dissemination of the information was therefore unlawful, and that the trial court properly ordered it stricken from the record and sealed.198 The language of the opinion was ambiguous as to whether the sealing pertained to the record in the case at bar or to all future public access.199 Although the Washington State Supreme Court granted review of Beltran, the parties settled, and the higher court never heard the case.200

In sum, Division I’s interpretation of the CRPA could prevent access to investigative information.201 Further application of the CRPA in this manner could make all investigative information resulting in a nonconviction inaccessible to the public, not just the criminal history record information about the incident.202 Although the Washington State Supreme Court did not review the cases establishing this precedent,203 the court has shown the willingness to reconsider Division I’s interpretation of the CRPA.204

197. Id. at 259, 989 P.2d at 612.
198. Id.
199. See id. (“Beltran was not authorized to obtain the information and the trial court properly remedied the unlawful dissemination of the information by ordering it to be stricken from the record and sealed.”). Division I has applied the CRPA to records beyond the statute’s definition of criminal history record information in two unpublished opinions since Beltran. State v. Soliz (In re Detention of Soliz), No. 44127-2-I, 2000 WL 965007, at *2, *9 (Wash. Ct. App. Div. 1 July 3, 2000); In re Detention of Thorell, No. 42237-5-I, 2000 WL 222815, at *7 (Wash. Ct. App. Div. 1 Feb. 22, 2000). In both cases, however, the court held disclosure to be permissible based on another statute requiring the Department of Corrections to disclose certain information about sexually violent predators to prosecutors. See Soliz, 2000 WL 965007, at *2, *9 (holding that release of police reports, Department of Corrections files, and mental health reports to psychologist was permissible under Washington Revised Code § 10.97.050(4)); Thorell, 2000 WL 222815, at *7 (holding that the same statutory provision permitted dissemination of Department of Corrections file). Since Beltran, the court has not considered the CRPA in a case where only the PDA supports disclosure.
201. See Beltran, 98 Wash. App. at 259–60, 989 P.2d at 611–12; see also supra notes 194–99 and accompanying text.
202. See Beltran, 98 Wash. App. at 259–60, 989 P.2d at 611–12; see also supra notes 194–99 and accompanying text.
203. Beltran, 140 Wash. 2d at 1021, 10 P.3d at 405 (dismissing upon settlement); Hudgens v. City of Renton, 110 Wash. 2d 1014 (1988) (denying review).
Disclosure of Nonconviction Information

V. THE WASHINGTON STATE SUPREME COURT SHOULD CURTAIL DIVISION I’S APPLICATION OF THE CRPA

Division I of the Washington State Court of Appeals has inappropriately extended the reach of the CRPA to entire bodies of investigative information.205 The Washington State Supreme Court could more appropriately reconcile the PDA and CRPA by redacting the narrow area of criminal history record information protected by the CRPA from investigative information that is otherwise publicly available under the PDA.206 This interpretation is supported by the text of the CRPA, which indicates that protected criminal history record information consists strictly of the three components of a rap sheet.207 Further, the federal regulations that mandated the CRPA are explicitly inapplicable to investigative information. The intended construction of the CRPA is to regulate the disclosure of rap sheet information only.208 Division I’s interpretation is also inconsistent with the Washington State Supreme Court’s narrowing of the privacy of investigative information under the PDA.209 The higher court has drawn a distinction between information that is private under the CRPA and information available under the PDA; this distinction is consistent with redacting CRPA-protected information from otherwise public records.210

A. Division I’s Broad Use of the CRPA To Address Entire Records of Nonconviction Investigations Conflicts with the Plain Text of the Statute

By applying the CRPA to all investigative information, Division I has expanded the statute’s narrow definition of criminal history record information. While Division I originally allowed the public to view investigative information held to be otherwise protected by the CRPA,211 the court has since indicated that the CRPA may prevent all access to

205. See supra Part IV.
206. See infra Part V.D.
207. See infra Part V.A.
208. See infra Part V.B.
209. See infra Part V.C.
210. See infra Part V.D.
investigative records about a nonconviction incident.\textsuperscript{212} The plain text of the CRPA does not protect investigative records, but instead protects a narrowly defined category of criminal history record information.\textsuperscript{213} The CRPA explicitly defines criminal history record information as the combination of three pieces of data: (1) identifiable descriptions of a person, (2) notations of arrests or formal criminal charges, and (3) the dispositions of those charges.\textsuperscript{214} Only when this information relates to a nonconviction incident does it have the potential to be private under the CRPA.\textsuperscript{215} The statute does not protect information outside this definition.\textsuperscript{216}

The CRPA’s “Inspection of Information by Subject” provision further demonstrates the statute’s distinction between criminal history record information and other investigative information.\textsuperscript{217} This provision explicitly grants individuals the right to review their own criminal history record information.\textsuperscript{218} However, the provision also indicates that the personal right of access extends only to criminal history record information, and not to investigative files.\textsuperscript{219} By distinguishing protected criminal history record information from unprotected investigative information, the CRPA indicates that investigative data falling outside of the three-part definition of criminal history record information is not addressed by the statute.\textsuperscript{220}

In line with such textual indications, state officials and commentators have interpreted the CRPA as inapplicable to investigative information.\textsuperscript{221} The Washington Attorney General’s Office \textit{Open Records & Open Meetings} deskbook specifies that “[i]nvestigative information does not fall within the definition of ‘criminal history record information.’”\textsuperscript{222} The deskbook further specifies that the release of

\begin{itemize}
  \item \textsuperscript{212} See Beltran v. State Dep’t of Soc. & Health Servs., 98 Wash. App. 245, 259, 989 P.2d 604, 611–12 (1999), dismissed upon settlement, 140 Wash. 2d 1021, 10 P.3d 405 (2000).
  \item \textsuperscript{213} \textsc{Wash. Rev. Code} § 10.97.030(1) (2000).
  \item \textsuperscript{214} \textit{Id.}
  \item \textsuperscript{215} \textit{Id.} § 10.97.050.
  \item \textsuperscript{216} \textit{See id.}
  \item \textsuperscript{217} \textit{See id.} § 10.97.080.
  \item \textsuperscript{218} \textit{Id.}
  \item \textsuperscript{219} \textit{Id.}
  \item \textsuperscript{220} \textit{See id.}
  \item \textsuperscript{221} See \textsc{Wash. State Office of the Attorney Gen.}, supra note 107, § 5.3; Horst, supra note 106, at 5-3 n.1.
  \item \textsuperscript{222} \textsc{Wash. State Office of the Attorney Gen.}, supra note 107, § 5.3.
\end{itemize}
Disclosure of Nonconviction Information

Police investigative information is governed by the PDA. Information from the library of Washington’s Municipal Research & Services Center indicates that although agencies apply the CRPA on a case-by-case basis, they commonly interpret the statute to cover criminal histories only. This predominant interpretation reflects the CRPA’s plain definition of criminal history record information as inclusive only of rap sheet information rather than entire investigative files. In applying the CRPA to protect investigative information from disclosure, Division I therefore broadens the scope of the CRPA’s text beyond common usage and understanding.

In addition, Division I’s interpretation of the CRPA undermines the PDA. The PDA states that courts are to promote its policy of disclosure by narrowly construing its exemptions. As the Washington State Supreme Court has indicated, courts should also follow this policy in interpreting exemptions passed more recently than the PDA. In Hearst Corp. v. Hoppe, the court held that a law passed after the PDA, which independently exempted “confidential income data” from disclosure, supplemented and defined the PDA’s existing right of privacy for taxpayers. This interpretation reconciled the two statutes and permitted the Hearst court to construe them as part of a harmonious scheme, as mandated by the principle of statutory construction on which the court relied. The Hearst court established that it is appropriate to reconcile the PDA with its exemptions by broadly construing the PDA’s disclosure policy and narrowly construing its exemptions.

In a manner similar to the income data exemption addressed in Hearst, the CRPA clarifies an existing exemption by explicitly designating certain types of private information within the context of investigative records. Addressing only the narrow category of criminal history record information, the CRPA is not a significant departure

223. Id.
224. See Horst, supra note 106, at 5-3 n.1.
225. See WASH. REV. CODE § 42.17.340(3).
228. See id.
229. See id.
230. Id. at 138, 580 P.2d at 254.
231. See WASH. REV. CODE § 10.97.050.
232. Id. § 42.17.310(1)(d).
233. Id. § 10.97.030(1).
from the public disclosure mandate of the PDA. Courts should therefore interpret the CRPA as a later exemption to the PDA. Thus, courts should read the CRPA in concert with the PDA’s “strongly worded mandate for broad disclosure of public records.” This interpretation reconciles and gives effect to both statutes, in line with the \textit{Hearst} decision.

In sum, Division I is not interpreting the CRPA in accord with its intended construction. The statute’s exemption for nonconviction criminal history record information is a specific and supplemental rule that defines the PDA’s privacy right in investigative records. When Division I applies the CRPA not only to criminal history record information, but also to entire investigative files that contain that data, it employs an inappropriately broad reading that contravenes the Washington State Supreme Court’s approach of reconciling the PDA and its later exemptions.

B. \textit{The Federal Regulations Mandating Washington’s CRPA Provide Guidance for Judicial Interpretation of the Statute}

The federal regulations that served as a model for the CRPA also indicate that Division I has misapplied the statute. The Washington State Supreme Court has shown its willingness to look to federal disclosure law by using interpretations of the federal Freedom of Information Act for guidance in construing the PDA. The text of the CRPA closely parallels the U.S. Department of Justice regulations that mandate it. The Department of Justice intended those regulations to protect the privacy of the rap sheet information shared between law enforcement agencies, not investigative information that is the work product of law enforcement officers.

\begin{enumerate}
  \item 234. \textit{Hearst}, 90 Wash. 2d at 127, 580 P.2d at 249.
  \item 235. \textit{See id. at 138, 580 P.2d at 254–55.}
  \item 236. \textit{WASH. REV. CODE § 10.97.030(2).}
  \item 238. \textit{Hearst}, 90 Wash. 2d at 139, 580 P.2d at 254–55.
  \item 239. \textit{Id. at 128, 580 P.2d at 249.}
  \item 241. \textit{See 28 C.F.R. §§ 20.3(i)–(p); 20.33; see also id. pt. 20 app.} (discussing the definition of criminal history record information contained in 28 C.F.R. § 20.3(d)).
\end{enumerate}
Disclosure of Nonconviction Information

The drafters of the federal regulations specified in their appendix that investigative information, or all other information contained in criminal justice agency reports, is not criminal history record information, and is thus beyond the scope of the model regulations. In text essentially identical to the federal regulations, Washington’s CRPA also defines criminal history record information to consist of the three basic components of a rap sheet. By not changing the language of the federal model law, Washington legislators implicitly adopted the appendix that came with the federal rules. Further, the legislative history behind the statute indicates that the state merely responded to the federal mandate, and did not expand upon it. If the legislature had an independent, broader interest in concealing investigative information, that interest is reflected neither in the text of the statute nor in the statute’s legislative history.

Division I’s application of the CRPA to investigative information is similarly misaligned with the purpose of the federal mandate as shown in additional legislative history. The federal mandate that inspired the CRPA was a response to the increasing use of computers by criminal justice agencies. As an incentive to those states that passed the requested legislation, the federal government promised funds for local record-keeping systems used to share information between agencies. This indicates that the chief federal concern was the accuracy and privacy of the rap sheet information shared electronically in response to requests for criminal histories. Rap sheets only include names paired with notations of arrests and their dispositions—the criminal history record information the CRPA purports to regulate.

The U.S. Supreme Court’s analysis in United States Department of Justice v. Reporters Committee for Freedom of the Press provides further guidance as to the difference between rap sheets and

242. 28 C.F.R. §§ 20.3(l)–(p), 20.33; see also id. pt. 20 app. (discussing the definition of criminal history record information contained in 28 C.F.R. § 20.3(d)).

243. See WASH. REV. CODE § 10.97.030(1).

244. STAFF OF SENATE JUDICIARY COMM., supra note 95, at 1; see also text accompanying supra note 95.

245. DALTON, supra note 88, at 1.

246. Id.

247. Id.

248. Id.; see also Criminal Justice Information Systems, 28 C.F.R. § 20.3(l) (2002).

investigative information. The decision confirmed that very different privacy considerations affect the release of rap sheets and the release of investigative information. The Court stated that rap sheets contain accumulated private information that does not reflect on the conduct of government. Thus, the combined elements of criminal history record information that would appear on a rap sheet are different from records such as police reports. Investigative reports are generally public under federal disclosure law, even when those reports contain information that could be compiled onto a rap sheet. Including investigative information within the CRPA’s coverage similarly conceals from public view not only criminal history record information about an individual, but also the work product of law enforcement officers. This interpretation expands the effect of the CRPA beyond individual privacy to the conduct of public agencies and the behavior of their employees—precisely the information the PDA was intended to bring into public view.

The Washington State Supreme Court has looked to interpretations of federal public records law to understand the meaning of Washington’s PDA. The court should likewise look to the U.S. Supreme Court’s analysis of rap sheet information and prevent Division I from subverting the PDA’s purpose with an overbroad application of the CRPA. The federal judicial interpretation, as well as the federal legislative history of the model law, demonstrates the distinction between the privacy of rap sheet information and the disclosure of investigative information. This distinction supports the use of the CRPA to redact rap sheet information from investigative information that is otherwise public.

251. Id.
252. Id. at 773.
253. See id. at 764.
254. See WASH. REV. CODE § 42.17.020(36) (applying to “any writing containing information relating to the conduct of government or the performance of any governmental or proprietary function”); Reporters Comm. for Freedom of the Press, 489 U.S. at 773.
255. WASH. REV. CODE § 42.17.260(1).
Disclosure of Nonconviction Information

C. Cowles Publishing Co. v. Spokane Police Department Supports Access to Investigative Records over Privacy

Division I’s narrowing of the public’s right to access investigative information is contrary to the Washington State Supreme Court’s holding in *Cowles Publishing Co. v. Spokane Police Department*, which expanded the circumstances when agencies must release that information.257 Although the CRPA was not a factor in the case, *Cowles* broadened the public’s right to access investigative records under the PDA.258 In limiting the *Newman* prohibition on disclosure of all open and active case records,259 the *Cowles* court rejected the idea that the defendant possessed an absolute right to privacy in the particulars of his arrest after his case had been referred to the prosecutor.260

The *Cowles* decision established a presumption that all information is subject to disclosure once a case is submitted for a prosecution decision, even if the information may later become nonconviction information protected by the CRPA.261 The court also presented a rationale for this result, noting that the public is generally aware of the presumption of innocence before adjudication.262 By establishing that investigative information is not generally private before it becomes nonconviction information,263 the *Cowles* court indicated that the PDA provides no reason why it should become private after a failure to convict.264 Therefore, the narrower right to privacy within the CRPA265 should not be applied to contravene the PDA by preventing disclosure of such investigative information across the board.

---

258. *Id.* at 479–80, 987 P.2d at 623–24.
259. *Id.* at 477–79, 987 P.2d at 622–23.
260. *Id.*
261. *Id.* at 481, 987 P.2d at 624.
262. *Id.* at 479, 987 P.2d at 623.
263. *Id.* at 481, 987 P.2d at 624.
264. *See id.*
265. *See supra* Part V.A–B.
D. The Washington State Supreme Court Can Best Reconcile the CRPA and PDA by Redacting Protected Criminal History Record Information from Investigative Records That Are Otherwise Public

Redacting information protected by the CRPA from information that is public under the PDA satisfies both statutes. The PDA requires redaction when private information appears in an otherwise public record.266 This approach is consistent with Washington State Supreme Court precedent that acknowledges a distinction between the information protected by the two statutes.267 Redaction is also the solution that best applies the two statutes as a harmonious scheme, in accord with the Washington State Supreme Court’s approach of reconciling the PDA’s mandate for disclosure with later exemptions to the law.268

The text of the PDA supports the solution of redacting protected nonconviction criminal history record information from public files, permitting access to the remaining information.269 The presence of a piece of private information in an otherwise public record is not a justification for withholding the entire record270 in the manner that Division I has suggested that entire investigative records may be withheld.271 Rather, the PDA requires that officials redact private information and release the remainder of the record.272 The CRPA governs only the disclosure of criminal history record information.273 The statute prohibits the disclosure of criminal history record information only when it concerns a nonconviction.274 When nonconviction criminal history record information appears in investigative records that are public under the PDA, agencies should therefore redact it. Under this approach, the information protected by the CRPA remains private, while remaining investigative information is available to the public.

266. WASH. REV. CODE § 42.17.260(1) (2000).
269. See WASH. REV. CODE § 42.17.260(1).
270. See id.
272. WASH. REV. CODE § 42.17.260(1).
273. See id. § 10.97.050.
274. See id.
Disclosure of Nonconviction Information

Redaction is consistent with the Washington State Supreme Court’s acknowledgment that the CRPA protects a smaller measure of police information than the PDA does. In Barfield, the court affirmed the release of investigative records from which “criminal records” had been deleted under the CRPA. As the Limstrom case progressed through the court system, opinions consistently referred to the possibility of some of the requested information receiving protection under the CRPA. This shows recognition by the court that protected data could be removed from files that would otherwise have to be disclosed. The court, however, has not clarified what information is protected by the CRPA. The text and history of the CRPA suggest that the statute addresses only criminal history record information—the three components that appear on a rap sheet Therefore, redacting this information from otherwise public records when it concerns a nonconviction would reconcile the CRPA and PDA in a logical extension of Washington State Supreme Court precedent.

Redaction also treats the PDA and CRPA as part of a single statutory scheme. The Washington State Supreme Court has indicated that the PDA and other laws that create exemptions to it should be reconciled into a harmonious scheme whenever possible. The court accomplishes this goal by reading the PDA’s broad policy of disclosure into later-created exemptions. Redaction of information protected by the CRPA would enable the disclosure of the underlying investigative information—consistent with the PDA’s policy and the court’s conclusion in Hearst Corp. v. Hoppe. At the same time, redaction would protect the nonconviction criminal history record information that the CRPA exempts from disclosure.

276. Barfield, 100 Wash. 2d at 885, 676 P.2d at 442.
278. See Limstrom, 136 Wash. 2d at 616, 963 P.2d at 880; Barfield, 100 Wash. 2d at 885, 676 P.2d at 442.
279. See supra Part V.A–B.
281. Id.
282. See WASH. REV. CODE § 42.17.340(3) (2000); Hearst, 90 Wash. 2d at 139, 580 P.2d at 255.
283. See WASH. REV. CODE § 10.97.050.
VI. CONCLUSION

The Washington State Supreme Court should end Division I’s overbroad application of the CRPA, and instead hold that the CRPA requires redaction of nonconviction criminal history record information from otherwise public documents. Redaction is consistent with the CRPA’s definition of criminal history record information and the PDA’s mandate that exemptions shall be interpreted narrowly. Neither the federal nor the state legislative history of the CRPA indicates an intention to include investigative information within the statute’s exemptions to the PDA. The Barfield and Limstrom decisions suggest a better approach to the CRPA by separating protected criminal history record information from documents not otherwise exempt from disclosure. The Washington State Supreme Court should therefore act to end Division I’s incorrect application of the CRPA that exempts investigative information from disclosure. Except for the redaction of nonconviction criminal history record information, courts should not apply the CRPA to investigative files.
CAPSIZED BY THE CONSTITUTION: CAN WASHINGTON STATE FERRIES MEET FEDERAL SCREENING REQUIREMENTS AND STILL PASS STATE CONSTITUTIONAL MUSTER?

David J. Perkins

Abstract: In response to the threat of international terrorism, the United States Coast Guard has issued new regulations requiring every ferry operator to begin screening passengers and vehicles for dangerous items. These new regulations will force Washington State Ferries (WSF), the agency responsible for operating the State’s ferry system, to either begin screening passengers and vehicles or face a possible shutdown. Compliance, however, poses problems for WSF because of privacy protections under the Washington State Constitution. Article I, section 7 of the state constitution contains an explicit protection of privacy. This section provides broader protections from warrantless searches by state actors than similar protections found in the Fourth Amendment of the United States Constitution. Apparently recognizing that article I, section 7 would not permit physical searches of vehicle interiors and trunks, WSF has proposed an alternative screening plan that would rely on non-intrusive technology or dogs to screen vehicles. This Comment argues that the use of these alternative screening methods will also be an unconstitutional, warrantless search under article I, section 7. Such searches will not satisfy any of the exceptions to this provision’s warrant requirement, and will make the searches unconstitutional intrusions into the private affairs of the ferry passengers. Despite its best efforts to find a compromise, WSF may still be faced with the choice between non-compliance with the federal regulations and violating its passengers’ constitutionally protected privacy rights.

Washington State Ferries (WSF) operates the largest ferry system in the United States.1 WSF describes itself as a “maritime highway” that carries 20,000 passengers each day.2 The ferry system consists of twenty-nine vessels and twenty ferry terminals.3 Additionally, WSF is the only practical means of transportation—and the sole means of commercial transport—between the mainland and populations living on Vashon Island and the San Juan Islands.4

WSF now faces a dilemma. New maritime security regulations5 issued by the U.S. Coast Guard as part of the nation’s “multi-front war

1. WASH. STATE FERRIES, AN INTRODUCTION TO THE LARGEST FERRY SYSTEM IN THE NATION 2 (2003).
2. Id. at 2.
3. Id. at 1.
4. Id. at 2.
against global terrorism\(^6\) require WSF to begin screening passengers and vehicles for dangerous items by June 30, 2004.\(^7\) All commercial ferry operators must develop a security plan that includes “a reasonable examination” of passengers and vehicles to ensure that no “dangerous substances and devices”\(^8\) are present.\(^8\) Any operator that fails to comply by the June 30, 2004 deadline faces the possible shutdown of its ferry operations.\(^9\)

Complying with these requirements, however, presents significant legal problems for WSF. Absent a warrant or an exception to the warrant requirement,\(^10\) opening and inspecting vehicle interiors and trunks would violate article I, section 7 of the Washington State Constitution.\(^11\) Article I, section 7’s explicit protection of privacy is broader than the similar protections of the Fourth Amendment\(^12\) of the U.S. Constitution.\(^13\) Under article I, section 7, warrantless searches are unconstitutional absent special circumstances.\(^14\)

Noting the protections of the state constitution, WSF has proposed an alternative, less intrusive screening plan to meet federal security requirements.\(^15\) Instead of opening vehicle interiors for inspection, this plan relies on the use of non-intrusive technology or dogs to screen vehicles.\(^16\) However, despite the less intrusive nature of this alternative plan, these methods will result in the same constitutional problems as opening and inspecting each vehicle compartment.

This Comment argues that WSF’s proposed screening plan violates article I, section 7 of the Washington State Constitution. As WSF may have recognized, physical inspections of vehicle trunks and interiors would constitute searches under the state constitution, triggering the protections of article I, section 7. However, WSF’s proposed use of non-intrusive technology or dogs will similarly constitute a search under

---

7. 33 C.F.R. § 104.115.
8. Id. § 101.105.
12. U.S. CONST. amend. IV (“The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated.”).
14. Id.
16. Id. at 3.
Capsized by the Constitution

Washington law. These searches will not satisfy any of the “jealously and carefully drawn” exceptions to the state constitution’s warrant requirement under article I, section 7. Despite WSF’s efforts to find a constitutionally acceptable method of screening vehicles, its current proposal will not satisfy either the requirements of the new federal maritime regulations or the protections of the state constitution.

Part I of this Comment provides an overview of the privacy protections of article I, section 7 of the Washington State Constitution, as well as the specific protections afforded to automobiles. Part II discusses the factors used by the Supreme Court of Washington to determine whether the use of technology or dogs is a search under article I, section 7. Part III describes the new federal maritime regulations and WSF’s proposal to comply with these regulations. Finally, Part IV argues that WSF’s use of non-intrusive technology or dogs to screen vehicles will constitute a search and will trigger the protections of article I, section 7. Part IV further argues that these searches will not satisfy any of the exceptions to the warrant requirement and will be an unconstitutional intrusion into the passengers’ private affairs. Part V concludes that a constitutional challenge to WSF’s plan under article I, section 7 will likely succeed, unless the court creates a new exception to the warrant requirement for minimally intrusive searches where the threat from terrorists is severe.

I. CITIZENS HAVE A PRESUMPTIVE RIGHT TO BE FREE OF WARRANTLESS SEARCHES UNDER ARTICLE I, SECTION 7

Article I, section 7 of the Washington State Constitution explicitly provides for a right to privacy that protects individuals from warrantless searches of their persons, homes, and vehicles. The provision states: “No person shall be disturbed in his private affairs, or his home invaded, without authority of law.” The Supreme Court of Washington has interpreted this section to create a body of jurisprudence independent from the Fourth Amendment to the U.S. Constitution. The protections of article I, section 7 are broader than those of the Fourth Amendment; the state protections prohibit warrantless searches unless justified by one

---

20. Hendrickson, 129 Wash. 2d at 70 n.1, 917 P.2d at 567 n.1.
of the recognized exceptions to the warrant requirement and include a historically recognized expectation of privacy in motor vehicles.22

A. Article I, Section 7 Provides a Broader Right to Privacy Than the Similar Protections Afforded by the Fourth Amendment

The Supreme Court of Washington has interpreted the text and history of the Washington State Constitution’s article I, section 7 as affording a broader right to privacy than that protected by the Fourth Amendment to the U.S. Constitution.23 In State v. Gunwall,24 the Supreme Court of Washington adopted six nonexclusive factors for courts to consider in determining whether a provision of the state constitution affords more protections than its federal counterpart.25 Applying those factors to article I, section 7, the court held that in certain circumstances this section provides broader protections than the Fourth Amendment.26 The court concluded that the framers of the state constitution intended to provide an explicit protection of privacy, not merely protection from unlawful searches and seizures, as evidenced by the significant textual differences between the state and federal provisions and the legislative history of article I, section 7.27

In the years immediately following the Gunwall decision, Washington courts would not address article I, section 7 contentions unless the parties adequately briefed the Gunwall factors.28 This adequate briefing requirement led to a number of cases in which Washington courts only addressed rights under the Fourth Amendment.29 Despite this initially inconsistent application, the court now considers it a matter of well-

21. Id. at 70, 917 P.2d at 568.
22. Id. at 70 n.1, 917 P.2d at 567 n.1.
25. These six factors are: (1) the textual language of the Washington State Constitution; (2) significant differences in the texts of parallel provisions of the federal and state constitutions; (3) state constitutional and common law history; (4) preexisting state law; (5) differences in the structure between the federal and state constitutions; and (6) matters of particular state interest or local concern. Id. at 61–62, 720 P.2d at 812–13.
26. See id. at 63–67, 720 P.2d at 814–15; see also infra Part I.D (providing examples of situations where article I, section 7 provides broader protections than the Fourth Amendment).
settled law that article I, section 7 provides broader protections than the Fourth Amendment.\textsuperscript{30}

B. \textit{Warrantless Searches Performed by State Actors Are Per Se Unreasonable and Violate Article I, Section 7}

An alleged violation of article I, section 7 must satisfy two requirements. First a “search” must occur.\textsuperscript{31} The Supreme Court of Washington has defined a search as an “intru[sion] into a person’s ‘private affairs.’”\textsuperscript{32} Under article I, section 7, the court defines the boundaries of “private affairs” by a traditional, objective standard, not the individual’s subjective perception of the intrusion.\textsuperscript{33} Specifically, the right to privacy includes “those privacy interests which citizens of this state have held, and should be entitled to hold, safe from governmental trespass absent a warrant.”\textsuperscript{34} This definition contrasts with the Fourth Amendment standard, which recognizes searches only in situations where individuals have both a reasonable and subjective expectation of privacy.\textsuperscript{35}

Second, the search must be performed by a state actor.\textsuperscript{36} Normally, for the purposes of article I, section 7, a state actor is a law-enforcement officer.\textsuperscript{37} However, other state employees acting in their official capacities can also be deemed state actors.\textsuperscript{38} For example, Washington courts have held that the actions of a tax appraiser,\textsuperscript{39} a city building inspector,\textsuperscript{40} a public utility district’s treasurer-comptroller,\textsuperscript{41} and school

\begin{itemize}
\item \textsuperscript{31} State v. Young, 123 Wash. 2d 173, 181, 867 P.2d 593, 597 (1994).
\item \textsuperscript{32} State v. Boland, 115 Wash. 2d 571, 577, 800 P.2d 1112, 1115 (1990) (quoting State v. Myrick, 102 Wash.2d 506, 510, 688 P.2d 151, 154 (1984)).
\item \textsuperscript{33} See Young, 123 Wash. 2d at 181, 867 P.2d at 597.
\item \textsuperscript{34} Id. (quoting State v. Myrick, 102 Wash. 2d 506, 511, 688 P.2d 151, 154 (1984)) (internal quotations omitted).
\item \textsuperscript{35} See id. (noting differences between private affairs inquiry under article I, section 7 and the reasonable expectation of privacy under the Fourth Amendment).
\item \textsuperscript{36} In re Maxfield, 133 Wash. 2d 332, 337, 945 P.2d 196, 198 (1997).
\item \textsuperscript{37} See id.
\item \textsuperscript{38} Id.
\item \textsuperscript{40} See City of Seattle v. McCready, 123 Wash. 2d 260, 271, 868 P.2d 134, 139–40 (1994).
\item \textsuperscript{41} See Maxfield, 133 Wash. 2d at 337, 945 P.2d at 199.
\end{itemize}
officials\textsuperscript{42} are searches by state actors that trigger the protections of article I, section 7.\textsuperscript{43}

Absent one of the six recognized exceptions,\textsuperscript{44} warrantless searches performed by state actors are “per se unreasonable” under article I, section 7.\textsuperscript{45} The Supreme Court of Washington first applied the per se unreasonable rule to warrantless searches approximately thirty years ago.\textsuperscript{46} Under this rule, warrantless searches are presumed to be impermissible, and the prosecution bears the burden of showing that these searches fall within an exception to the warrant requirement.\textsuperscript{47}

C. Washington Courts Permit Warrantless Searches Under Six Exceptions to the Warrant Requirement

The Supreme Court of Washington recognizes six “jealously and carefully drawn” exceptions to the general rule that warrantless searches are per se unreasonable under article I, section 7.\textsuperscript{48} These six exceptions are: (1) searches incident to a valid arrest; (2) inventory searches; (3) plain view; (4) investigative stops; (5) consent; and (6) exigent circumstances.\textsuperscript{49} Under article I, section 7, the court does not recognize an exception for circumstances where a threat to public safety may outweigh individual privacy rights.\textsuperscript{50}

Washington case law has further defined and limited these exceptions to the warrant requirement. A search incident to arrest only applies during the arrest process and is generally limited to searches for weapons or destructible evidence.\textsuperscript{51} Inventory searches are limited to

\begin{itemize}
\item \textsuperscript{42} See Kuehn v. Renton Sch. Dist. No. 403, 103 Wash. 2d 594, 602, 694 P.2d 1078, 1083 (1985).
\item \textsuperscript{43} Maxfield, 133 Wash. 2d at 337, 945 P.2d at 199.
\item \textsuperscript{44} See infra Part I.C.
\item \textsuperscript{45} State v. Hendrickson, 129 Wash. 2d 61, 71, 917 P.2d 563, 568 (1996).
\item \textsuperscript{46} Id. at 70–71, 917 P.2d at 568.
\item \textsuperscript{47} Id. at 70, 917 P.2d at 568.
\item \textsuperscript{48} Id. Interpreting the Fourth Amendment, the Supreme Court of Washington has applied a “community caretaking” exception to allow evidence obtained by police encounters involving emergency aid and routine checks on health and safety where these encounters were completely divorced from any criminal investigation. However, the court did not address if these actions violated article I, section 7 because the parties had not properly briefed the Gunwall factors. State v. Kinzy, 141 Wash. 2d 373, 385 n.33, 5 P.3d 668, 675 n.33 (2000). The court has not recognized the community caretaking exception in the article I, section 7 jurisprudence.
\item \textsuperscript{49} Hendrickson, 129 Wash. 2d at 71, 917 P.2d at 568.
\item \textsuperscript{50} See infra Part I.C.3.
\end{itemize}
instances where officers need to precisely catalog property taken into
custody to defend against claims of theft and vandalism. The plain
view exception only allows for a warrantless visual search where
officers “(1) have a prior justification for the intrusion; (2) inadvertently
discover the incriminating evidence; and (3) immediately recognize the
item as contraband.” The investigative stop is an exception that allows
law-enforcement officers to briefly detain persons for questioning if the
officers reasonably suspect these persons of criminal activity. The
remaining two exceptions, consent and exigent circumstances, require a
more detailed explanation.

1. Consent Must Be Given Voluntarily To Satisfy the Consent
   Exception

   Coerced or involuntary consent does not satisfy the consent exception
to the warrant requirement. Washington courts have struck down
consent searches where the method employed to obtain consent was
coercive. For example, in State v. Ferrier, the Supreme Court of
Washington struck down “knock and talk” procedures in which heavily
armed law-enforcement officers entered a suspect’s home and asked for
consent to search. The court held that this “show of force” was
“inherently coercive,” and any subsequently obtained consent was
invalid.

   Additionally, the Supreme Court of Washington has determined that
requiring patrons to consent to searches as a condition of admission to a
public event constitutes impermissible coercion. In Jacobsen v. City of
Seattle, the court found coercion where admission to a rock concert
was conditioned on consent to pat down searches. The court rejected

54. See Terry v. Ohio, 392 U.S. 1, 24 (1968).
56. See Ferrier, 136 Wash. 2d at 115, 960 P.2d at 933.
57. 136 Wash. 2d at 118–19, 960 P.2d at 934.
58. Id. at 118–19, 960 P.2d at 934.
59. Id. at 115, 960 P.2d at 933.
62. Id. at 670, 658 P.2d at 654–55.
the notion that these searches were consensual despite video evidence showing concert goers willfully and good-naturedly complying with the search requirement.\(^{63}\) Where admission is conditioned on consent, this consent is coerced and does not satisfy the consent exception to the warrant requirement.\(^{64}\)

Under article I, section 7, the Supreme Court of Washington has allowed the notion of implied consent in narrow circumstances.\(^{65}\) In State v. Curran,\(^{66}\) the court upheld a statute that declared a driver to be “deemed to have given consent” to a blood alcohol test\(^{67}\) if the driver has been arrested and the arresting officer has “reasonable grounds to believe the [driver] had been driving . . . under the influence of intoxicating liquor.”\(^{68}\) The court first determined that taking the arrested driver’s blood and testing it for evidence of intoxication was a search.\(^{69}\) The court then held that the search was constitutional under a theory of implied consent because there “was a clear indication that [the blood test] would reveal evidence of [the driver’s] intoxication.”\(^{70}\) The court concluded that this “clear indication” was met by evidence showing that the driver smelled of alcohol and appeared disoriented with watery, bloodshot eyes.\(^{71}\) Thus, although the theory of implied consent appears to broaden the consent exception, the court has only upheld this theory in situations where a law-enforcement officer had a “clear indication” that the search of an arrested suspect would uncover evidence of a crime.

2. The Exigent Circumstances Exception Requires Both Probable Cause and the Impracticality of Obtaining a Warrant

Another exception to the warrant requirement, exigent circumstances, exists when law-enforcement officers can show both probable cause and that obtaining a warrant is impractical because of an immediate need to search for and seize evidence.\(^{72}\) First, the officers must have probable

\(^{63}\) Id. at 674, 658 P.2d at 656–67.
\(^{64}\) See id. at 674, 658 P.2d at 656.
\(^{67}\) WASH. REV. CODE § 46.20.308 (2003).
\(^{68}\) Curran, 116 Wash. 2d at 179, 804 P.2d at 561.
\(^{69}\) Id. at 184, 804 P.2d at 564.
\(^{70}\) Id.
\(^{71}\) Id.

cause that the items to be seized are connected with criminal activity and will be found in the place to be searched. Probable cause is an individualized suspicion, or nexus, between the suspects to be searched and the criminal activity. For example, in State v. Counts, law-enforcement officers used a tracking dog to follow a fleeing suspect to his house. The Supreme Court of Washington concluded that officers had sufficient probable cause to search the home based on an individualized suspicion that they would find the suspect within.

Second, the officers must show that special circumstances made it impractical for them to obtain a warrant. The court has found “special circumstances” in situations such as the pursuit of a fleeing suspect or where evidence may be destroyed. If the state can show both probable cause and the impracticality of obtaining a warrant, then a warrantless search is permissible under the exigent circumstances exception.


Under article I, section 7, the Supreme Court of Washington does not recognize an exception to the warrant requirement for situations in which the need to curb a threat to public safety can justify warrantless searches. Federal courts have upheld searches at airports and

73. Id. at 377.
76. Id. at 59, 659 P.2d at 1089.
77. See id. at 59–60, 659 P.2d at 1089–90 (recognizing that the officers had sufficient probable cause to obtain a warrant, but invalidating the warrantless search because the officers were not in “hot pursuit”).
78. Johnson, supra note 72, at 503.
79. See Counts, 99 Wash. 2d at 60, 659 P.2d at 1090 (recognizing that hot pursuit could create exigent circumstances, but holding that hot pursuit did not exist when suspect fled into home under the observation of law-enforcement officers).
81. Johnson, supra note 72, at 503.
82. See United States v. Skipwith, 482 F.2d 1272, 1276 (5th Cir. 1973) (holding that searches at airports did not violate the Fourth Amendment where the danger to the public was severe and the metal detector searches used were minimally intrusive).
courthouses under the Fourth Amendment by balancing public security, the efficacy of the search, and the degree of intrusion involved. Prior to Gunwall, in the infancy of independent article I, section 7 jurisprudence, the Supreme Court of Washington included “airport and courthouse searches” within its list of exceptions to the warrant requirement but never applied this exception to a specific case. Following Gunwall, however, the court has made consistent reference to only the six exceptions described above when deciding a case under article I, section 7.

D. Washington Courts Recognize a Legitimate Expectation of Privacy in Motor Vehicles and Refuse To Permit Warrantless Searches Absent an Exception to the Warrant Requirement

There is a long-standing history in Washington of recognizing the privacy interests of individuals in their automobiles. As the Supreme Court of Washington said in City of Seattle v. Mesiani, “[f]rom the earliest days of the automobile in this state, this court has acknowledged the privacy interest of individuals and objects in automobiles.” More explicitly, the court declared in State v. Stroud that “a person in possession of a vehicle has a legitimate expectation of privacy under article I, section 7 in the vehicle . . . [and] articles within the vehicle which . . . are not visible.” Accordingly, under Washington law, warrantless searches of motor vehicles, like searches of persons or homes, are per se unreasonable under article I, section 7. The Supreme Court of Washington has struck down several types of warrantless motor

83. See Downing v. Kunzig, 454 F.2d 1230, 1233 (6th Cir. 1972) (holding that searches at courthouses did not violate the Fourth Amendment).
85. See id. at 672, 658 P.2d at 655.
86. See State v. Hendrickson, 129 Wash. 2d 61, 71, 917 P.2d 563, 568 (1996); see also City of Seattle v. Mesiani, 110 Wash. 2d 454, 456–58, 755 P.2d 775, 776–77 (1988) (refusing to create an exception for a situation where sobriety checkpoints were used to curb the public safety danger posed by drunk driving).
87. See Mesiani, 110 Wash. 2d at 457, 755 P.2d at 777.
89. Id. at 456–57, 755 P.2d at 777.
90. 106 Wash. 2d 144, 720 P.2d 436 (1986).
91. Id. at 152, 720 P.2d at 441.
vehicle searches under article I, section 7, including the warrantless search of locked vehicle compartments after the arrest of the owner in *Stroud*93 and the use of sobriety checkpoints in *Mesiani*.94

These rulings starkly contrast with U.S. Supreme Court Fourth Amendment jurisprudence involving searches and seizures of motor vehicles. Under the Fourth Amendment, the U.S. Supreme Court allows law-enforcement officers to search the passenger compartment of a vehicle, and any locked compartments within, after the arrest of the driver.95 Taking a more restrictive stance, the *Stroud* court held that exigent circumstances, such as the possibility that evidence could be destroyed or the vehicle could be driven away, justified a search incident to arrest of the passenger compartment, but did not justify a search of locked vehicle compartments under article I, section 7.96 The court noted that locking a vehicle compartment produced “additional privacy expectations” that were “objectively justifiable.”97

Similarly, while the Supreme Court of Washington prohibits sobriety checkpoints under article I, section 7, these checkpoints are permissible under the Fourth Amendment.98 The U.S. Supreme Court allowed sobriety checkpoints under the Fourth Amendment, reasoning that the public safety benefits of curbing drunk driving outweighed the “slight” intrusion into the rights of law-abiding motorists.99 In contrast, the Supreme Court of Washington struck down sobriety checkpoints under article I, section 7 in *Mesiani*.100 Despite evidence that cataloged the “slaughter on our highways,”101 the *Mesiani* court refused to justify these intrusions based on “an inference from statistics that there are inebriated drivers in the area.”102 Instead of adopting a balancing test that weighed the public safety benefit against the intrusion into private affairs, the

---

93. See *Stroud*, 106 Wash. 2d at 152, 720 P.2d at 441.
96. *Stroud*, 106 Wash. 2d at 152, 720 P.2d at 441.
97. Id.
court limited its inquiry to the traditional exceptions to the warrant requirement.\textsuperscript{103}

Although one Washington State Court of Appeals decision did uphold the warrantless search of all vehicles on a ferry, the court based the holding on the Fourth Amendment, not article I, section 7, and limited it to specific factual circumstances.\textsuperscript{104} In State v. Silvernail,\textsuperscript{105} Division I of the Washington State Court of Appeals upheld the search of all vehicles on a ferry.\textsuperscript{106} As two suspects fled the scene of an armed robbery on Vashon Island, a victim overheard a portion of the suspects’ conversation and inferred that they were heading for the Seattle ferry.\textsuperscript{107} When the ferry arrived in Seattle, police officers began to search each vehicle on the ferry and discovered weapons and “suspected narcotics” in Silvernail’s vehicle.\textsuperscript{108} Relying on Fourth Amendment precedent, the Washington State Court of Appeals upheld the search.\textsuperscript{109} Although the Supreme Court of Washington did not review the case, the high court has subsequently noted the limited scope of the Silvernail opinion. The court declined to extend the Silvernail reasoning in Mesiani and commented that the Silvernail decision was “expressly limited to situations in which there was reliable information that a serious felony had recently been committed.”\textsuperscript{110}

In sum, article I, section 7 provides broader protections from warrantless searches of persons and vehicles than the Fourth Amendment. The article I, section 7 protections include a general prohibition on warrantless searches that is limited by six judicially recognized exceptions. Consent can serve as an exception to the warrant requirement if it is voluntary. Similarly, exigent circumstances can qualify as an exception if the law-enforcement officer performing the search can show that sufficient probable cause existed and that obtaining a warrant was impractical. Washington courts do not recognize an

\textsuperscript{103} See id. at 456–57, 755 P.2d at 776–77.
\textsuperscript{104} Id. at 458 n.1, 755 P.2d at 777 n.1.
\textsuperscript{105} 25 Wash. App. 185, 605 P.2d 1279 (1980).
\textsuperscript{106} Id. at 191, 605 P.2d at 1283.
\textsuperscript{107} Id. at 186–88, 605 P.2d at 1281.
\textsuperscript{108} Id. Silvernail was not one of the two armed robbery suspects. He had the misfortune of being on the wrong ferry at the wrong time.
\textsuperscript{109} Id. at 190–91, 605 P.2d at 1282–83 (noting the danger or unrestrained roadblock searches, but holding that the search was permissible because of the minimal intrusion of the search and its “reasonable likelihood of success”).
exception to the warrant requirement where the public safety benefit outweighs an intrusion into private affairs. Finally, the broader protections of article I, section 7 include a historic recognition of an individual’s legitimate expectation of privacy in a motor vehicle.

II. UNDER ARTICLE I, SECTION 7, THE USE OF TECHNOLOGICAL DEVICES OR DOGS MAY BE A SEARCH

Washington courts limit the ability of law-enforcement officers to use technology or dogs to intrude into private affairs. Noting that technology “races ahead with ever increasing speed,” eroding expectations of privacy “without our awareness, much less our consent,” Washington courts continue to protect traditional privacy rights under article I, section 7.111 In certain situations, the courts classify the use of technology or dogs as a search subject to the warrant requirement of article I, section 7.112 This classification depends on the vantage point, the level of intrusiveness, and the constitutional protections involved in the use.113

A. The Use of Infrared Devices To Detect Activity Within a Home Constitutes a Search

The Supreme Court of Washington has declared that the use of infrared detection devices to observe activities within a home constitutes a search under article I, section 7, and must therefore satisfy one of the exceptions to the warrant requirement.114 In State v. Young,115 law-enforcement officers used infrared thermal detection devices to detect excessive heat emanating from a residence suspected of containing a marijuana growing operation.116 To determine whether the use of this technology constituted a search under article I, section 7, the court considered three factors: the lawfulness of the vantage point, the intrusiveness of the means used, and the nature of the property

111. State v. Young, 123 Wash. 2d 173, 184, 867 P.2d 593, 598 (1994).
113. Young, 123 Wash. 2d at 182–84, 867 P.2d at 597–98.
114. See id. at 184, 867 P.2d at 598–99.
116. Id. at 177–78, 867 P.2d at 595.
targeted. Although the law-enforcement officer in *Young* was observing from a lawful vantage point, the court noted that the intrusiveness of the infrared detector allowed officers to “see through the walls” and that the target of the observation was a home. Based on these three factors, the court ruled that the use of technology was an intrusion into both a home and private affairs and was therefore a search under article I, section 7. Because the officers had not obtained a warrant or satisfied an exception to the warrant requirement, the court invalidated this procedure as an unreasonable search.

**B. A Dog Sniff May Be a Search if It Intrudes into a Constitutionally Protected Place**

Under Washington law, a dog sniff may constitute a search depending on the object sniffed and circumstances surrounding the sniff. State appellate court decisions have held that a dog sniff does not constitute a search under article I, section 7 if the defendant has no reasonable expectation of privacy in the object being sniffed and the dog sniff is minimally intrusive. In *State v. Boyce*, the court ruled that a dog sniff around the defendant’s safety deposit box was not a search under article I, section 7. The *Boyce* court reasoned that the defendant had no expectation of privacy in the bank vault and that it was only minimally intrusive for an officer invited into the bank to use the dog in this manner. The court followed nearly identical reasoning in *State v. Stanphill*, and ruled that a dog sniff of a package at a post office was also not a search. In *State v. Wolohan*, the court held that a dog sniff of a package at a bus terminal was not a search but expressed “grave
Capsized by the Constitution

doubts” about extending this ruling to a dog sniff of personal effects on or near a person. 130 This situational approach contrasts with the U.S. Supreme Court’s blanket ruling that dog sniffs never constitute a search under the Fourth Amendment.131

Although the Supreme Court of Washington has not ruled on these appellate court dog sniff rulings, 132 it has stated that a dog sniff might constitute a search under article I, section 7 in some situations. 133 The court made this statement in Young as part of its discussion of the use of infrared detectors. 134 Without elaborating, the court said that dog sniffs targeted at an object or location that is “subject to heightened constitutional protection” might be searches under article I, section 7. 135

The Washington State Court of Appeals subsequently applied the Young factors to determine whether a dog sniff outside of a garage was a search under article I, section 7. 136 Relying on the dicta in Young concerning dog searches, 137 the court looked at the lawfulness of the vantage point, the intrusiveness of the means used, and the nature of the property targeted to determine whether a dog sniff was a search. 138 The court stated that the vantage point of the sniff was lawful, but the dog sniff was intrusive because it allowed officers to “see through the walls.” 139 Furthermore, the target of the sniff, a garage attached to a home, was subject to higher constitutional protections. 140 The court concluded that the dog sniff of the garage was a search and was unreasonable because the officers had not obtained a warrant or satisfied any of the exceptions to the warrant requirement. 141

130. Id. at 820 n.5, 598 P.2d at 425 n.5. The court applied the Fourth Amendment because the search took place in Arizona. Id. at 814, 598 P.2d at 422.
131. United States v. Place, 462 U.S. 696, 707 (1983) (holding that a dog sniff of luggage is not a search because it does not require the opening of the luggage, does not expose non-contraband items to public view, and only discloses the presence or absence of contraband, and therefore is limited in both the manner of intrusion and the content of the information revealed).
132. Neither Boyce, Stanphill, nor Wolohan were appealed to the Supreme Court of Washington.
134. Id.
135. Id.
137. See Young, 123 Wash. 2d at 187–88, 867 P.2d at 600–01.
139. Id. at 635, 962 P.2d at 853–54.
140. Id.
141. See id. at 635–37, 962 P.2d at 853–54.
In sum, under Washington law, the privacy protections of article I, section 7 limit the use of technology and dogs to intrude into private affairs. Determining whether the use of a technological device by law-enforcement officers is a search depends on the lawfulness of the vantage point, the intrusiveness of the means used, and the nature of the property targeted. Applying similar factors, the use of dogs may also constitute a search under article I, section 7.

III. WSF PLANS TO USE NON-INTRUSIVE TECHNOLOGY AND DOGS TO SCREEN VEHICLES

WSF is proposing an alternative security program, using non-intrusive technology or dogs, to meet the new security requirements outlined in the National Maritime Security Initiative (NMSI). The NMSI regulations require all ferry operators to search a variable percentage of passengers and vehicles prior to boarding. Perhaps recognizing that the privacy protections of article I, section 7 of the Washington State Constitution would prevent state agents from conducting searches of passengers and vehicles, WSF has submitted an alternative security plan.

A. The NMSI Regulations Require All Ferry Operators To Search a Percentage of Passengers and Their Motor Vehicles and To Deny Passage to Passengers Who Refuse To Consent to These Searches

Pursuant to its power under the Commerce Clause, Congress enacted the Maritime Transportation Security Act of 2002 (MTSA) to

142. WASH. STATE FERRIES, supra note 15, at 1.
144. Id. § 104.265.
146. Since Gibbons v. Ogden, 22 U.S. (9 Wheat.) 1 (1824), the U.S. Supreme Court has recognized the authority of Congress to regulate navigable waterways and ports derived from the Commerce Clause of the United States Constitution. U.S. CONST. art. I, § 8, cl. 3. If federal and state regulations of maritime activities conflict, the Court asks whether Congress has left the states any room to impose regulations of their own, not whether Congress has the authority to regulate in this area. See United States v. Locke, 529 U.S. 89, 108 (2000). Congress cannot compel states to enforce or implement federal programs. See Prinz v. United States, 521 U.S. 898, 935 (1997) (holding that Congress could not require local law-enforcement officers to conduct background checks as part of a federal program regulating the interstate sale of firearms even though Congress had the authority under the Commerce Clause to regulate the sales); New York v. United States, 505 U.S. 144, 188 (1992) (holding that Congress could not compel New York to either regulate low-
Capsized by the Constitution

safeguard the nation’s ports and waterways.\textsuperscript{148} International terrorists have already demonstrated their inclination to use the nation’s maritime transportation system to further their attacks.\textsuperscript{149} Ahmed Ressam, an Algerian national, was arrested for smuggling bomb making materials from Victoria, British Columbia, to Port Angeles, Washington, aboard a WSF ferry in the spare tire compartment of his rental car.\textsuperscript{150} Recognizing that other terrorists could use the United States’ maritime transportation systems to further their goals, the MTSA aims to protect the nation’s ports from terrorist attacks and to prevent their use by terrorists.\textsuperscript{151}

To implement the MTSA, the U.S. Coast Guard issued new regulations in the form of the NMSI.\textsuperscript{152} These regulations condition continued vessel operation on each vessel operator’s submission and implementation of an acceptable Vessel Security Plan.\textsuperscript{153} Any vessel operator failing to comply by June 30, 2004\textsuperscript{154} must cease operations.\textsuperscript{155} An acceptable security plan must provide for the “reasonable examination” of a randomly selected percentage of passengers and

\textsuperscript{147} 46 U.S.C. §§ 70102–70103 (Partial Revision 2003).

\textsuperscript{148} U.S. DEP’T OF HOMELAND SEC., PROTECTING AMERICA’S PORTS 3 (2003).

\textsuperscript{149} Id.


\textsuperscript{151} U.S. DEP’T OF HOMELAND SEC., supra note 148, at 3.

\textsuperscript{152} Id.


\textsuperscript{154} Id. § 104.115(b).

\textsuperscript{155} 46 U.S.C. § 70103(c)(5) (Partial Revision 2003).
vehicles for “dangerous substances and devices.”156 This percentage will increase as the Marine Security level increases.157 At the highest security level, operators must screen all passengers and vehicles158 even though the U.S. Coast Guard may not be able to identify a certain vessel as a specific target.159 Additionally, all vessel operators must post signs stating that “[b]oarding the vessel is deemed valid consent to screening or inspection,” and “[f]ailure to consent or submit to screening or inspection will result in denial or revocation of authorization to board.”160

B. WSF Has Proposed an Alternative Security Program To Satisfy NMSI Requirements

WSF’s proposed alternative security program will vary significantly from the standard screening procedures outlined in the NMSI.161 A provision of the NMSI allows vessel operators to submit alternative security programs that satisfy the same screening requirements discussed above but that deviate in their specific screening procedures.162 WSF submitted an alternative security plan that will use “non-intrusive technology” and “explosive detection dogs” to screen vehicles.163 In a press release explaining this plan, WSF indicated that opening and

156. 33 C.F.R. § 104.265(c)(1). The NMSI defines “screening” as “a reasonable examination of persons, cargo, vehicles, or baggage . . . to ensure that dangerous substances and devices, or other items that pose a real danger of violence or a threat to security are not present.” Id. § 101.105. “Dangerous substances or devices” are defined as “any material, substance, or item that may cause damage or injury to any person, vessel, facility, harbor, port, or waters” and is “unlawful to possess under applicable Federal, State or local law.” Id.

157. Id. § 104.265. The Marine Security (MARSEC) level is “set to reflect the prevailing threat environment to the marine elements of the national transportation system.” Id. § 101.105. MARSEC-1, the lowest level of security, is defined as “the level for which minimum appropriate protective security measures shall be maintained at all times.” Id. MARSEC-2 is “the level for which appropriate additional protective security measures shall be maintained for a period of time as a result of heightened risk of a transportation security incident.” Id. MARSEC-3 is the “level for which further specific protective security measures shall be maintained for a limited period of time when a transportation security incident is probable or imminent, although it may not be possible to identify the specific target.” Id.

158. Id. § 104.265(g)(1).

159. Id. § 101.105.

160. Id. § 104.265(e)(2).

161. WASH. STATE FERRIES, supra note 15, at 1–3.

162. 33 C.F.R. § 104.140.

163. WASH. STATE FERRIES, supra note 15, at 3. WSF has not specified what types of “non-intrusive technology” it will be using. See id.
Capsized by the Constitution

inspecting vehicle interiors and trunks might lead to constitutional challenges. The proposed alternative plan is WSF’s attempt to satisfy both the federal regulatory and state constitutional requirements. Assuming that the U.S. Coast Guard approves WSF’s alternative plan, these new security measures will go into effect by June 30, 2004.

IV. WSF’S PROPOSED PROGRAM WILL BE AN UNCONSTITUTIONAL SEARCH

Despite WSF’s attempt to craft an alternate security plan that complies with article I, section 7, the use of technology and dogs to screen vehicles will still violate the privacy rights of ferry passengers. The use of non-intrusive technology and dogs to ascertain the contents of vehicles at ferry terminals will constitute a warrantless search under article I, section 7. These searches will not satisfy any of the exceptions to the warrant requirement, and will therefore be impermissible under article I, section 7.

A. The Use of Non-Intrusive Technology or Dogs To Screen Vehicles Will Be a Search Subject to the Warrant Requirement of Article I, Section 7

The opening and inspection of a vehicle’s interior and trunk by a state actor constitutes a warrantless search triggering the protections of article I, section 7. WSF’s proposed alternative, the use of technology or dogs to screen vehicles, will also constitute a search. Using technology or dogs, like all other searches by state actors, will be unreasonable per se and unconstitutional under article I, section 7, absent a warrant or one of the exceptions to the warrant requirement.

165. Id.
166. WASH. STATE FERRIES, supra note 15, at 3.
167. See infra Part IV.A.
168. See infra Part IV.B.
169. See infra Part IV.A.1.
170. See infra Part IV.A.2–3.
171. See supra Part I.A–C.
1. The Opening and Inspecting of Vehicle Interiors and Trunks by State Employees Are Searches Under Article I, Section 7

As WSF may have suspected, the opening and inspection of vehicle interiors and trunks by WSF employees satisfies both requirements of a search under article I, section 7. First, the Supreme Court of Washington has held that opening and inspecting the contents of locked vehicle compartments is an intrusion into the private affairs of the vehicle owner. Second, the WSF employees performing these screenings are state employees acting in their official capacities. Like other state employees, such as a tax appraiser, a city building inspector, a public utility district’s treasurer-comptroller, and school officials, Washington courts will find WSF employees to be state actors as well.

2. The Use of Non-Intrusive Technology To Screen Vehicles Will Also Constitute a Search Under Article I, Section 7

WSF’s plan to use non-intrusive technology to screen vehicles will still qualify as a search under article I, section 7. The Supreme Court of Washington considered three factors in ruling that the use of infrared detection equipment to search houses for suspected marijuana growing operations constituted a search under article I, section 7: the lawfulness of the vantage point, the intrusiveness of the means used, and the nature of the property targeted. The court concluded that although the officers were using the device from a lawful vantage point, the action still constituted a search because the intrusiveness of the device enabled the officers to “see through the walls” to observe private affairs and violated the constitutional protection of the home.

Under the same analysis, the use of non-intrusive technology to screen vehicles at a ferry terminal will also be a search. The screening at the ferry terminals will be from a lawful vantage point, but WSF will use this technology to “see through” the locked compartments of the vehicles. A vehicle does not have the same level of constitutional protection as a home, but Washington courts have nonetheless

173. See supra notes 32–43 and accompanying text.
175. See supra notes 38–43 and accompanying text.
177. Id. at 183, 867 P.2d at 598.
recognized a strong expectation of privacy in vehicles. Under article I, section 7, the Supreme Court of Washington has frequently noted the history of a legitimate, and constitutionally protected, expectation of privacy in both vehicles and objects hidden inside. This expectation of privacy increases when the owner locks items within vehicle compartments. Applying the Young factors, the use of technology from a lawful vantage point to “see into” the compartments of a constitutionally protected vehicle, although less obtrusive than opening and inspecting vehicle compartments, will still be a search under article I, section 7.

3. The Use of Dogs To Screen Vehicles Will Also Constitute a Search Under Article I, Section 7

Applying the Young factors to dog sniffs, the use of dogs to detect explosives or contraband in motor vehicles will also qualify as a search under article I, section 7. Relying on dicta in Young that a dog sniff of an area “subject to heightened constitutional protections” might constitute a search under article I, section 7, Division I of the Washington State Court of Appeals ruled in State v. Dearman that a dog sniff of a garage was a search. The Dearman court applied the Young factors and determined that the dog sniff was from a lawful vantage point outside the garage, but that the sniff allowed the officers to “see through the walls” of the garage, and the garage was constitutionally protected because of its proximity to a home. The court concluded that this dog sniff satisfied the Young factors and was therefore a search.

Similarly, WSF’s proposed use of dogs to screen motor vehicles will constitute a search and trigger the protections of article I, section 7. Like the dog sniff in the Dearman case, WSF’s proposed dog sniff will be from a lawful vantage point. However, these sniffs will allow WSF

178. Stroud, 106 Wash. 2d at 152, 720 P.2d at 441.
180. See Stroud, 106 Wash. 2d at 152–53, 720 P.2d at 441.
181. Young, 123 Wash. 2d at 188, 867 P.2d at 600.
183. Id. at 635, 962 P.2d at 853.
184. Id.
185. Id.
employees to see inside vehicles to “ensure that dangerous substances and devices . . . are not present.” 186 WSF’s use of dogs will produce the same kind of intrusion that existed in the Dearman case.

Finally, Washington courts have continuously recognized a legitimate expectation of privacy in vehicles under article I, section 7. 187 This legitimate expectation distinguishes dog sniffs of vehicles from the previous appellate court decisions holding that dog sniffs were not searches where the owner had no expectation of privacy in the object or location sniffed. 188 Motor vehicles do not share the highest protections afforded to homes. 189 However, the Young court only referred to “heightened constitutional protection” in its discussion of when a dog sniff would be a search, 190 and vehicles should meet this standard. The Supreme Court of Washington has consistently held that article I, section 7 provides a higher level of constitutional protection for vehicles than the Fourth Amendment. 191 The court has noted that these protections increase when the driver locks a vehicle compartment. 192 Thus, the use of dogs to screen motor vehicles will satisfy the Young factors and will constitute a search under article I, section 7.

B. Non-Intrusive Technology and Dog Searches Would Not Satisfy Any Recognized Exception to the Warrant Requirements

Having established that the use of technological devices or dogs will constitute searches under article I, section 7, these searches are per se unreasonable unless they satisfy one of the recognized exceptions to the warrant requirement. 193 This Comment assumes that four of the six exceptions—search incident to arrest, inventory searches, plain view, and investigative stops—would be inapplicable to routine screenings at ferry terminals. Because WSF is not arresting its passengers, or impounding the passengers’ vehicles, these screenings would not be

188. See supra notes 122–129 and accompanying text.
193. See Hendrickson, 129 Wash. 2d at 71, 917 P.2d at 568.
Capsized by the Constitution

searches incident to arrest\textsuperscript{194} or inventory searches.\textsuperscript{195} WSF intends to screen for items inside closed vehicle containers—not just inadvertently discover objects in plain view.\textsuperscript{196} Finally, WSF will not have the reasonable and particularized suspicion necessary to conduct an investigative stop.\textsuperscript{197} If neither of the two remaining exceptions—consent and exigent circumstances—are satisfied, the proposed searches will be unconstitutional under article I, section 7.\textsuperscript{198}

1. Conditioning Ferry Passage on Consent to Non-Intrusive Technology and Dog Screenings Will Not Satisfy the Consent Exception to the Warrant Requirement

WSF will not be able to satisfy the consent exception by requiring passengers to consent to searches as a condition of passage on the ferry.\textsuperscript{199} The consent exception requires voluntarily,\textsuperscript{200} uncoerced\textsuperscript{201} consent. Conditioning ferry travel on consent to search is analogous to the rock concert searches struck down by the Supreme Court of Washington in \textit{Jacobsen v. City of Seattle}.\textsuperscript{202} In \textit{Jacobsen}, concert patrons were required to submit to searches as a condition of their admission to the concert.\textsuperscript{203} The court rejected the notion that the rock concert searches were consensual despite video evidence showing

\textsuperscript{194} A search incident to arrest must be conducted “[d]uring the arrest process, including the time immediately subsequent to the suspect’s being arrested, handcuffed, and placed in a patrol car” and is generally limited to searches for weapons or destructible evidence. \textit{Stroud}, 106 Wash. 2d at 152, 720 P.2d at 441.

\textsuperscript{195} The rationale for allowing inventory searches is to allow law-enforcement officers to precisely catalog property to defend against claims of theft, vandalism, or negligence. \textit{See State v. Dugas}, 109 Wash. App. 592, 597, 36 P.3d 577, 580 (2001).

\textsuperscript{196} The plain view exception will justify a warrantless search where an officer: (1) has a prior justification for the intrusion; (2) inadvertently discovers the incriminating evidence; and (3) immediately recognizes the item as contraband. \textit{See State v. Myers}, 117 Wash. 2d 332, 346, 815 P.2d 761, 769 (1991).

\textsuperscript{197} Investigative stops allow law-enforcement officers to briefly detain persons for questioning if the law-enforcement officer reasonably suspects these persons of criminal activity. \textit{See Terry v. Ohio}, 392 U.S. 1, 24 (1968).

\textsuperscript{198} \textit{See supra} Part I.C; notes 48–49 and accompanying text.


\textsuperscript{200} \textit{See supra} note 55 and accompanying text.

\textsuperscript{201} \textit{See supra} note 56 and accompanying text.


\textsuperscript{203} \textit{See id. at 669–70, 658 P.2d at 654}.
patrons cheerfully complying with the search requirement.\(^{204}\) In contrast to the activity at issue in \textit{Jacobsen}, use of the ferry system is a necessity for 20,000 daily passengers, especially for residents of Vashon Island and the San Juan Islands who commute by ferry between their homes and work places.\(^{205}\) As in \textit{Jacobsen}, Washington courts are likely to find any consent of ferry passengers invalid because continued use of the ferry system is conditioned on giving this consent.

Additionally, WSF cannot rely on the “implied consent” doctrine to uphold these proposed searches. This limited extension of the consent requirement only applies to instances where there is clear evidence that the search will find evidence of a specific crime.\(^{206}\) In \textit{State v. Curran}, the Supreme Court of Washington upheld a statute that implied a driver’s consent to a blood alcohol test, but only where there “was a clear indication that it would reveal evidence of [the driver’s] intoxication.”\(^{207}\) WSF will not have a similar indication that the screening of vehicles will uncover evidence of any crime. Instead, WSF will be randomly selecting and screening a certain number of vehicles\(^{208}\) out of over 20,000 daily passengers.\(^{209}\) WSF will have no reason to believe that passengers in every screened vehicle are involved in criminal activity. Consequently, WSF will not be able to use the doctrine of “implied consent” to justify the non-intrusive technology and dog screening of vehicles at ferry terminals.

2. \textit{These Searches Will Not Satisfy the Exigent Circumstances Exception Because There Is No Individualized Probable Cause To Search Each Vehicle}

Without individualized probable cause to search each vehicle, the WSF’s proposal to use non-intrusive technology or dogs will not satisfy the exigent circumstances exception. Under this exception, the state actor conducting the search must have probable cause to believe that evidence of a crime, in this case terrorism, will be found in the vehicle being searched.\(^{210}\) To establish probable cause, there must be a nexus

\(^{204}\) \textit{Id.} at 674, 658 P.2d at 656–57.

\(^{205}\) \textit{WASH. STATE FERRIES, supra} note 1, at 1–2.


\(^{207}\) \textit{Id.} at 184, 804 P.2d at 564.

\(^{208}\) \textit{WASH. STATE FERRIES, supra} note 15, at 2.

\(^{209}\) \textit{WASH. STATE FERRIES, supra} note 1, at 1–2.

\(^{210}\) \textit{See supra} Part I.C.2.
between the vehicles searched and the crime of terrorism.211 The NMSI regulations will require WSF to search randomly selected vehicles—or all vehicles—depending on the threat level.212 Because these vehicles will not be selected based on any particularized information, they will not have the requisite nexus to terrorism. Without this nexus, no probable cause will exist sufficient to justify these searches under the exigent circumstances exception.

The Washington State Court of Appeals’ decision in State v. Silvernail upholding a search of all vehicles on a ferry213 does not undermine this conclusion. First, the court in Silvernail only considered the protections provided by the Fourth Amendment, not article I, section 7.214 Second, the Supreme Court of Washington limited the Silvernail decision to “situations in which there was reliable information that a felony had recently been committed,”215 such as when the fleeing suspect was believed to be on a particular ferry. A similar situation will not arise during the proposed screenings at ferry terminals. Even under the highest level of threat, the U.S. Coast Guard recognizes that it may not be able to identify a specific target.216 Therefore, during normal operations, WSF will not have the kind of “reliable information” to search vehicles on all of its ferries that the Silvernail court used to justify searches on a particular ferry.

3. Despite the Threat to Public Safety Posed by Terrorists, These Searches Will Not Be Permissible Under Article I, Section 7

WSF cannot use the threat of terrorism as an exception to the warrant requirement. Unlike Fourth Amendment decisions upholding searches at airports217 and courthouses,218 the Supreme Court of Washington does not recognize an exception to the warrant requirement under article I, section 7 for situations where a threat to public safety is sufficient to justify warrantless searches. The court refers to only the six “jealously

211. See supra Part I.C.2.
214. Id. This opinion was decided six years before State v. Gunwall, 106 Wash. 2d 54, 720 P.2d 808 (1986), and before Washington courts began to routinely address state constitutional issues.
216. See 33 C.F.R. § 101.105.
217. See United States v. Skipwith, 482 F.2d 1272, 1276 (5th Cir. 1973).
and carefully drawn” exceptions when deciding a case under article I, section 7.219 In Mesiani, the court had the opportunity to create a new exception to allow sobriety checkpoints and curb drunk driving.220 Although presented with evidence of “slaughter on our highways,”221 the court refused to draw “an inference from statistics that there are inebriated drivers in the area.”222 Before Gunwall, the court had acknowledged, but never applied, the balancing test used to justify searches at airports and courthouses where the danger to the public was severe and the intrusion was minimal.223 However, when confronted with the danger posed by drunk driving, the court rejected the notion of a balancing test exception under article I, section 7, whereby the public safety benefit could outweigh individual privacy rights.224 Following this precedent, the court is unlikely to “draw inferences” that there are terrorists on a particular ferry and unlikely to allow an inferred threat to public safety to outweigh real violations of privacy rights.

V. CONCLUSION

By June 30, 2004, WSF plans to begin screening vehicles through the use of non-intrusive technology or dogs to avoid a possible shutdown for failure to satisfy the NMSI security requirements. Although WSF is attempting to find a constitutionally and publicly acceptable alternative to physically opening and searching vehicles, ascertaining the contents of vehicles—whether by physical search or dog sniff—is still a search under article I, section 7. Because these searches do not fit any of the exceptions to the warrant requirement, the searches are likely to be unconstitutional.

A constitutional challenge to these screenings, although supported by current article I, section 7 precedent concerning dog sniffs and vehicle searches, will not necessarily be successful. When considering this challenge, Washington courts will be influenced by the changed security concerns following the terrorist attacks of September 11, 2001 and the disastrous possibility of a ferry system shutdown, neither of which were

220. See Mesiani, 110 Wash. 2d at 458, 755 P.2d at 777.
222. Mesiani, 110 Wash. 2d at 458 n.1, 755 P.2d at 777 n.1.
Capsized by the Constitution

present in prior article I, section 7 cases. Ultimately, the courts may modify article I, section 7 jurisprudence to add a new exception for minimally intrusive searches where the threat from terrorists is severe.
PROTESTORS HAVE FOURTH AMENDMENT RIGHTS, TOO: IN GRAVES V. CITY OF COEUR D’ALENE, THE NINTH CIRCUIT CLOUDS CLEARLY ESTABLISHED LAW GOVERNING SEARCHES

Holly Vance

Abstract: In Graves v. City of Coeur d’Alene, the United States Court of Appeals for the Ninth Circuit concluded that a police officer should not have arrested a protestor at an Aryan Nations parade when the protestor refused to allow the officer to search his backpack. The court held that the arrest was illegal because the officer had no probable cause to believe the protestor was carrying a weapon. However, the court also held that the arresting officer was entitled to qualified immunity and thus not liable for his violation of the protestor’s rights. Qualified immunity is a privilege that shields a public official from liability in situations where the underlying substantive law is not clearly established. In Graves, the court held that under the circumstances surrounding the Aryan Nations parade, the standard for probable cause to search the protestor’s backpack was not clearly established. This Note argues that the police officer in Graves searched the protestor without sufficient individualized suspicion. Instead, as a basis for the search the officer relied on a broad profile of otherwise ordinary conduct that would include a number of innocent individuals. Because the law was clearly established at the time of the search that police officers must have some individualized suspicion of a person to perform a search, and that individualized suspicion cannot be based on ordinary conduct that would include a number of innocent people, the defendant police officer was not entitled to qualified immunity.

In light of recent terrorist attacks, police officers have the critical task of maintaining public safety while protecting constitutionally guaranteed civil rights. The attacks of September 11, 2001, the Oklahoma City bombing, and the bombing at the Atlanta Olympics have caused widespread fear of terrorism. At the same time, police officers have responded with increased vigilance. However, there is a risk that increased security measures are at odds with this nation’s history of upholding civil liberties.

One of those liberties is the right to be free of unreasonable searches.\(^4\) That right is protected by the Fourth Amendment to the United States Constitution, which prohibits police from searching people’s possessions unless there is probable cause to believe that the possessions contain “contraband or evidence of a crime.”\(^5\) To establish probable cause, police officers must have individualized suspicion of a particular person.\(^6\) Individualized suspicion cannot be based on a broad profile of otherwise ordinary conduct that would include a number of innocent individuals.\(^7\)

The opinion of the U.S. Court of Appeals for the Ninth Circuit in *Graves v. City of Coeur d’Alene*\(^8\) illustrates the conflict between security concerns and Fourth Amendment rights. In *Graves*, a Coeur d’Alene police officer arrested a protestor at an Aryan Nations parade for obstructing the officer by refusing to allow the officer to search his bulky backpack.\(^9\) The *Graves* court held that the arrest was unlawful because the officer lacked probable cause for the search, and the protestor was within his rights to refuse the search.\(^10\) However, the court also concluded that the arresting officer was entitled to qualified immunity because the standard for probable cause for the search of a person’s backpack on a public street during a demonstration had not been clearly established.\(^11\)

This Note argues that the *Graves* court improperly granted the officer qualified immunity from civil liability because the probable cause standard governing the search of the demonstrator’s backpack was clearly established. Part I of this Note outlines U.S. Supreme Court and Ninth Circuit precedent that guide police officers’ determinations of probable cause and reasonable suspicion for searches and seizures. Part II reviews the doctrine of qualified immunity and examines the “clearly established” standard. Part III outlines the facts, procedural history, holding, and rationale of the *Graves* decision. Part IV argues that the Ninth Circuit erred in holding that the officer in the *Graves* case was

---

4. *See* U.S. CONST. amend. IV.
8. 339 F.3d 828 (9th Cir. 2003).
9. *Id.* at 836–37.
10. *Id.* at 844.
11. *Id.* at 847–48.
Fourth Amendment Rights of Protestors

entitled to qualified immunity because in making a probable cause determination, the officer relied on a broad profile of ordinary conduct that would implicate many innocent individuals, and clearly established law prohibited the use of such a broad profile as a basis for individualized suspicion in support of probable cause.

I. POLICE MUST HAVE INDIVIDUALIZED SUSPICION FOR SEARCH OR SEIZURE OF A PARTICULAR SUSPECT

The Fourth Amendment protects “[t]he right of the people to be secure in their . . . effects, against unreasonable searches and seizures.”

Police officers must have probable cause to search a person’s effects. Brief investigative stops (Terry stops) require reasonable suspicion—a lower level of suspicion than probable cause. To establish probable cause or reasonable suspicion, officers must have “individualized suspicion” that a person has broken the law, is breaking the law, or is about to break the law. Officers cannot show individualized suspicion through the use of a broad profile consisting of otherwise ordinary conduct that would implicate a number of innocent individuals. Exceptions to the Court’s requirement for individualized suspicion are limited to special needs programs and certain administrative searches.

A. Individualized Suspicion Is a Necessary Component of Probable Cause and Reasonable Suspicion Determinations

The Fourth Amendment prohibits police from conducting unreasonable searches or seizures. A search occurs when police invade an area, or look through a possession, in which someone has a reasonable expectation of privacy, while a seizure occurs when a

12. U.S. Const. amend. IV.
15. See City of Indianapolis v. Edmond, 531 U.S. 32, 37 (2000) (holding that individualized suspicion is necessary for a search or seizure to be reasonable under the Fourth Amendment except for limited situations involving special needs and administrative searches); Chandler v. Miller, 520 U.S. 305, 308, 313 (1997) (same).
17. See Edmond, 531 U.S. at 37.
18. U.S. Const. amend. IV.
person is restrained from moving by the use of physical force or submission to authority.20 Seizures include arrests and Terry stops.21 Police officers must have probable cause to conduct a legal search22 or arrest.23 To lawfully conduct a Terry stop, officers must have reasonable suspicion, which is a lower predicate than probable cause.24

An unreasonable search is a search based on less than probable cause.25 The U.S. Supreme Court has held that probable cause for a search exists when “there is a fair probability that contraband or evidence of a crime will be found in a particular place.”26 The Court has further defined probable cause as “‘less than evidence which would justify condemnation’ or conviction,”27 but more than “bare suspicion.”28 A Terry stop is unreasonable when it is grounded on less than reasonable suspicion.29 The reasonable suspicion standard is satisfied when police officers reasonably believe “that criminal activity may be afoot.”30

To conduct a lawful search or Terry stop, police officers must base their determination of probable cause or reasonable suspicion on individualized suspicion about a particular suspect.31 In making a determination of individualized suspicion for a search or Terry stop, the issue “is not whether particular conduct is ‘innocent’ or ‘guilty,’ but the degree of suspicion that attaches to particular types of noncriminal acts.”32 In Ybarra v. Illinois,33 the U.S. Supreme Court considered whether officers had probable cause to search Ventura Ybarra when, acting under a warrant that did not include Ybarra, the officers searched a tavern where he was a patron.34 Officers obtained the warrant based on

28. Id.
34. Id. at 88.
Fourth Amendment Rights of Protestors

a tip from an informant who reported seeing heroin packets in the possession of the bartender and having the heroin offered for sale to him by the bartender.\textsuperscript{35} During the search of the tavern, officers frisked Ybarra and removed a cigarette packet that was later found to contain heroin.\textsuperscript{36} The Court held that there was no probable cause for the officers’ search of Ybarra because they did not have any basis for individualized suspicion of him.\textsuperscript{37} Similarly, in \textit{Terry v. Ohio},\textsuperscript{38} the Court held that a brief investigative stop is permissible only when officers have a reasonable suspicion based on facts particular to the individual that a person is engaged in criminal activity.\textsuperscript{39}

Although police officers may consider the surrounding circumstances in their analysis of probable cause or reasonable suspicion, individualized suspicion of a particular suspect is always necessary for a search or \textit{Terry} stop.\textsuperscript{40} For example, in \textit{Ybarra}, the Court held that regardless of the suspicious circumstances at the tavern, including known drug trafficking by the bartender and some of the patrons, officers did not have probable cause to perform a search when they had no individualized suspicion of Ybarra.\textsuperscript{41} Likewise, in \textit{Brown v. Texas},\textsuperscript{42} the Court held that officers could not base reasonable suspicion entirely on the surrounding circumstances in the context of a \textit{Terry} stop.\textsuperscript{43} In \textit{Brown}, police officers stopped Zackary Brown after observing him walking away from another man in an alley that was frequently used for drug trafficking and demanded that he identify himself.\textsuperscript{44} The Court concluded that Brown’s stop was not lawful under the \textit{Terry} standard because officers did not base their determination of reasonable suspicion on any facts that were particular to Brown.\textsuperscript{45}

\textsuperscript{35} Id. at 87–88.
\textsuperscript{36} Id. at 88–89.
\textsuperscript{37} See id. at 90–92.
\textsuperscript{38} 392 U.S. 1 (1968).
\textsuperscript{39} See id. at 21.
\textsuperscript{41} See \textit{Ybarra}, 444 U.S. at 90–91.
\textsuperscript{42} 443 U.S. 47 (1979).
\textsuperscript{43} Id. at 52–53.
\textsuperscript{44} Id. at 48–49.
\textsuperscript{45} Id. at 52.
B. Individualized Suspicion Cannot Be Based on Characteristics That Would Implicate a Number of Innocent Individuals

In the context of a Terry stop, police officers may not base their determination of individualized suspicion of a person on broad profiles of otherwise ordinary conduct that would implicate a number of innocent individuals.46 In Reid v. Georgia,47 the U.S. Supreme Court held that an officer from the Drug Enforcement Administration impermissibly stopped the defendant without individualized suspicion of wrongdoing.48

The officer based the stop on the following factors, which he believed fit a typical “drug courier profile”:49 (1) the defendant had traveled from a city frequently implicated in drug trafficking; (2) he arrived early in the morning; (3) the defendant and his companion did not check any luggage; and (4) the defendant and his companion appeared to be trying to hide the fact they were traveling together.50 The Court reasoned that only the last of these factors related to the particular conduct of the defendant.51 It concluded that “[t]he other circumstances describe a very large category of presumably innocent travelers, who would be subject to virtually random seizures were the Court to conclude that as little foundation as there was in this case could justify a seizure.”52 While the Court concluded that the fourth factor did cast some suspicion on the defendant, it determined that the defendant’s conduct in relation to his traveling companion was insufficient to establish reasonable suspicion.53 The Reid Court therefore held that the officer did not have sufficient individualized suspicion of the defendant to support a Terry stop.54

In a similar case, the Ninth Circuit also concluded that officers cannot ground individualized suspicion on ordinary behavior that would include many innocent individuals.55 In United States v. Rodriguez,56 border

46. See Reid v. Georgia, 448 U.S. 438, 441 (1980); United States v. Rodriguez, 976 F.2d 592, 595–96 (9th Cir. 1992). There is no case law applying this principle to the probable cause standard for searches.
47. 448 U.S. 438 (1980).
48. See id. at 441.
49. Id. at 440–41.
50. Id.
51. See id. at 441.
52. Id.
53. See id.
54. See id.
55. See United States v. Rodriguez, 976 F.2d 592, 595–96 (9th Cir. 1992).
56. 976 F.2d 592 (9th Cir. 1992).
Fourth Amendment Rights of Protestors

control agents stopped Ramiro Rodriguez after observing him driving alone down the highway. When the agents decided to perform a *Terry* stop of Rodriguez, they relied on the following factors: (1) the road he was driving on was a “notorious route for alien smugglers”; (2) he did not acknowledge the agents as they drove by in a marked vehicle; (3) he was driving a car the agents believed could be used for alien smuggling; (4) he looked at the agents in the rearview mirror several times; (5) the car appeared to be “heavily loaded”; and (6) Rodriguez was a Hispanic male. The *Rodriguez* court held that these factors did not provide sufficient individualized suspicion of Rodriguez to justify his stop because each factor was otherwise normal behavior, and even taken together, the factors would ensnare a number of innocent people. It concluded that the court “must not accept what has come to appear to be a prefabricated or recycled profile of suspicious behavior very likely to sweep many ordinary citizens into a generality of suspicious appearance merely on hunch.”

While police officers cannot base individualized suspicion wholly on otherwise ordinary conduct that would include a number of law-abiding individuals, otherwise innocent factors, when grouped together, can sometimes amount to reasonable suspicion. In *United States v. Sokolow*, the U.S. Supreme Court held that officers had reasonable suspicion to stop a traveler at an airport when he: (1) paid $2100 in cash for two airline tickets from a roll of twenty-dollar bills; (2) traveled under an alias; (3) listed his original destination as Miami, a city known as a source of illegal drugs; (4) stayed in Miami for only forty-eight hours even though his round-trip flight took twenty hours; (5) appeared nervous; and (6) checked no luggage. The Court held that although none of these factors individually would have established reasonable suspicion, taken together, they amounted to reasonable suspicion. Thus, ordinary behavior, when examined by police as a whole, may amount to individualized suspicion. However, officers may not base

57. Id. at 593.
58. Id. at 594–95.
59. See id. at 595–96.
60. Id.
63. See id. at 3, 9–10.
64. Id. at 9–10.
65. See id.
individualized suspicion on a profile of ordinary conduct that would implicate many innocent individuals.\textsuperscript{66}

\textbf{C. Officers May Only Conduct a Search Unsupported by Individualized Suspicion as Part of a Special Needs Program or Administrative Search}

Police officers may only conduct a search without individualized suspicion of a crime in limited circumstances.\textsuperscript{67} The U.S. Supreme Court has upheld searches ungrounded in individualized suspicion only in cases where the searches were part of an organized special needs program or a specific administrative search.\textsuperscript{68} A special needs program must be based on standardized criteria and cannot be founded on evidence or suspicion of criminal conduct.\textsuperscript{69} Constitutionally permissible special needs programs include random drug testing of student athletes,\textsuperscript{70} drug and alcohol testing of certain railway employees,\textsuperscript{71} and road checkpoints to detect intoxicated drivers and illegal aliens.\textsuperscript{72} Examples of allowable suspicionless administrative searches include inspections of a “closely regulated” junkyard business to uncover evidence of car theft\textsuperscript{73} and inspections of buildings to determine the cause of a fire.\textsuperscript{74}

Special needs programs are constitutional only if the searches are aimed at addressing a valid state objective and are based on special needs, beyond those of normal law enforcement.\textsuperscript{75} For example, in \textit{Michigan Department of State Police v. Sitz},\textsuperscript{76} the Court held that a random highway checkpoint for alcohol intoxication was not unreasonable because the state had a valid interest, beyond the normal needs of law enforcement, in preventing drunk driving.\textsuperscript{77} However, the

\begin{itemize}
\item \textsuperscript{66} See Reid v. Georgia, 448 U.S. 438, 441 (1980); United States v. Rodriguez, 976 F.2d 592, 595–96 (9th Cir. 1992).
\item \textsuperscript{67} See City of Indianapolis v. Edmond, 531 U.S. 32, 37 (2000).
\item \textsuperscript{68} Id.
\item \textsuperscript{69} See Florida v. Wells, 495 U.S. 1, 3–4 (1990).
\item \textsuperscript{71} Skinner v. Ry. Labor Executives’ Ass’n, 489 U.S. 602, 634 (1989).
\item \textsuperscript{72} Edmond, 531 U.S. at 32.
\item \textsuperscript{73} New York v. Burger, 482 U.S. 691, 712 (1987).
\item \textsuperscript{74} Michigan v. Tyler, 436 U.S. 499, 509 (1978).
\item \textsuperscript{76} 496 U.S. 444 (1990).
\item \textsuperscript{77} See id. at 455.
\end{itemize}
Fourth Amendment Rights of Protestors

Court reached the opposite conclusion in *City of Indianapolis v. Edmond*, 78 when it held that random road checkpoints for illegal drugs were unconstitutional.79 The Court reasoned that because the purpose of the *Edmond* checkpoints was almost indistinguishable from general crime prevention, allowing such searches would have effectively circumvented the Court’s requirement for individualized suspicion for searches related to general crime prevention.80

In sum, when making a determination of probable cause or reasonable suspicion, police officers may consider the circumstances surrounding the situation in addition to the factors particular to the suspect.81 However, officers must have some individualized suspicion of a person to establish probable cause or reasonable suspicion.82 Individualized suspicion cannot be based on a profile consisting of otherwise ordinary conduct that is likely to implicate a number of innocent individuals.83 Absent individualized suspicion, officers may only conduct a search or seizure as part of a special needs program that advances a valid state objective extending beyond the goals of normal law enforcement or in conjunction with a valid administrative search.84

II. QUALIFIED IMMUNITY PROTECTS PUBLIC OFFICIALS FROM CIVIL LIABILITY IN LIMITED CIRCUMSTANCES

Qualified immunity allows public officials to avoid liability for violating a person’s civil liberties.85 The doctrine protects officials from liability in situations where the law was not clear at the time of the violation,86 thus ensuring “that before they are subjected to suit, officers are on notice their conduct is unlawful.”87 The U.S. Supreme Court established the doctrine of qualified immunity as a means of balancing

79. See id. at 48.
80. See id. at 44.
82. See *Edmond*, 531 U.S. at 37; Chandler v. Miller, 520 U.S. 305, 313 (1997).
83. See Reid v. Georgia, 448 U.S. 438, 441 (1980); United States v. Rodriguez, 976 F.2d 592, 595–96 (9th Cir. 1992).
the rights of citizens against the public interest in allowing public officials to function effectively.\textsuperscript{88}

In \textit{Saucier v. Katz},\textsuperscript{89} the Court announced a two-part test for determining whether an official is entitled to qualified immunity.\textsuperscript{90} First, a plaintiff must establish that an official violated one of the plaintiff’s constitutional rights.\textsuperscript{91} When considering whether plaintiff’s rights have been violated, a court must view the facts in the light most favorable to the plaintiff.\textsuperscript{92} The court then considers whether the right was “clearly established” at the time it was violated.\textsuperscript{93} A right is clearly established if it would be objectively clear to a reasonable official that the official’s conduct was unlawful.\textsuperscript{94}

When evaluating whether a right is clearly established, the U.S. Supreme Court has held that courts should look at precedent applying the legal rule in a similar factual situation.\textsuperscript{95} However, the test is whether case law makes a right “apparent,” not whether “the very action in question has previously been held unlawful.”\textsuperscript{96} For example, in \textit{Mendoza v. Block},\textsuperscript{97} police officers used a dog to track a robbery suspect.\textsuperscript{98} The suspect claimed that the officers’ use of the dog was an excessive use of force, but the officers raised a qualified immunity defense.\textsuperscript{99} The question before the Ninth Circuit was whether the law regarding use of a police dog was clearly established.\textsuperscript{100} The officers claimed that it was not, because there were few cases dealing with the issue of dogs.\textsuperscript{101} The \textit{Mendoza} court disagreed, finding that a number of cases had clearly established the law regarding use of force in general,\textsuperscript{102} and that a

\textsuperscript{89} 533 U.S. 194 (2001).
\textsuperscript{90} See id. at 201.
\textsuperscript{91} See id.; see also \textit{Hope}, 536 U.S. at 736.
\textsuperscript{92} See \textit{Katz}, 533 U.S. at 201.
\textsuperscript{93} Id.
\textsuperscript{95} Id.
\textsuperscript{96} Id.
\textsuperscript{97} 27 F.3d 1357 (9th Cir. 1994).
\textsuperscript{98} Id. at 1358–59.
\textsuperscript{99} Id. at 1359.
\textsuperscript{100} See id. at 1360–61.
\textsuperscript{101} See id. at 1361. The officers claimed there was no case law that addressed the use of police dogs, although the \textit{Mendoza} court noted that there were other decisions that discussed this issue. Id.
\textsuperscript{102} See id. at 1361–62.
Fourth Amendment Rights of Protestors

reasonable police officer would have known that those cases applied to any use of force, whether through “use of a baton, use of a gun, or use of a dog.” Thus, to clearly establish a legal principle, case law need not consider every factual scenario.

III. IN GRAVES THE NINTH CIRCUIT HELD THAT A POLICE OFFICER WAS ENTITLED TO QUALIFIED IMMUNITY

In Graves v. City of Coeur d’Alene, the Ninth Circuit considered whether a police officer violated the civil liberties of a protestor at an Aryan Nations parade by arresting him when he refused to consent to a search of his backpack. The court held that the officer did not have probable cause to search the backpack and the arrest was unlawful. However, the court concluded that the officer was entitled to qualified immunity from civil liability because the law had not been clearly established regarding the degree of weight that the officer could assign to the surrounding circumstances of the parade when making his determination of probable cause.

A. The Jury Found That Officer Dixon Did Not Violate Crowell’s Civil Rights

The dispute in Graves arose from events occurring at an Aryan Nations parade in Coeur d’Alene, Idaho, in 1998. Local police expected that several groups, including the Jewish Defense League, would protest the parade. The police were worried about the potential for violence at the parade, given the reputations of both the Aryan Nations and the Jewish Defense League, and the fact that the president of the Jewish Defense League had announced that Coeur d’Alene’s streets would “run red with blood.” Adding to the tension surrounding the parade was a report that someone had stolen ammonium nitrate, a chemical used in construction blasting, from a construction site.

103. Id. at 1362.
104. See id.
106. See id. at 845.
107. Id. at 847–48.
108. Id. at 834–37.
109. Id. at 834.
110. Id. (internal quotations omitted).
Because of their concerns about violence, police officers were searching the bags of many of the people protesting the march. One of the plaintiffs in *Graves*, Jonathan Crowell, attended the parade to voice his opposition to the Aryan Nations. Crowell walked along the parade route wearing a heavy backpack and carrying a sign that read, “Earth first, hatred last.” Officer Dixon, a member of the Coeur d’Alene Police Department, demanded that he be allowed to search Crowell’s backpack. Crowell refused, stating that it was his constitutional right not to consent to a search. Officer Dixon then demanded to search the backpack, but Crowell continued to refuse to consent to the search. Finally, Officer Dixon arrested Crowell for obstructing a police officer. The one other protestor who refused to allow the officer to search his bag was also arrested. After Crowell’s arrest, Officer Dixon searched Crowell’s backpack and discovered jars of peanut butter, jelly, and applesauce, in addition to bread, shoes, and clothes. In 1999, Crowell sued Officer Dixon in the U.S. District Court for the District of Idaho for false arrest under 42 U.S.C. § 1983. At the trial, Officer Dixon testified that while Crowell had definitely asserted that the search violated his civil rights, he had not verbally or physically assaulted the officer. Officer Dixon also testified that he suspected Crowell’s backpack contained explosives.

---

111. *Id.*
113. *Graves*, 339 F.3d at 834.
114. *Id.* at 835.
115. *Id.* at 836.
116. *Id.*
117. *Id.*
118. *Id.* at 837. Officer Dixon called his supervisor, Lieutenant Hotchkiss, by radio before arresting Crowell. Lieutenant Surplus overheard the call and replied, “[I]f he won’t let us look in the pack, [you] need[] to arrest him.” *Id.* (internal quotations omitted).
119. Brief for Appellant at 3, *Graves* (No. 02-35119).
120. *Graves*, 339 F.3d at 837. On April 14–15, 1999, Crowell was prosecuted for obstructing a police officer pursuant to Idaho Code § 18-705, which defines obstruction of an officer as conduct that “resists, delays, or obstructs any public officer, in the discharge, or attempt to discharge, of any duty of his office.” *Id.* at 833, 840 (quoting IDAHO CODE § 18-705). The trial resulted in a hung jury. *Id.* at 833. As of June 1, 2001, Crowell was awaiting retrial. *Id.* at 833 n.1.
121. *Id.* at 833. Five protestors, Jonathan Crowell, Gary Bizek, Lori Graves, Jeffrey Kerns, and Kenneth Malone, were originally plaintiffs in the case. *Id.* at 833 n.2. All claims except those made by Crowell and Bizek were dismissed on partial summary judgment. *Id.* Crowell also named Officer Surplus as a defendant on a theory of supervisory liability. *Id.*
122. *Id.* at 836–37.
Fourth Amendment Rights of Protestors

because it was “heavy” and looked like it contained “round cylindrical type objects.” However, he also admitted to searching the bags of approximately fourteen other protestors that day and could not remember targeting the bags of these protestors for any particular reason, such as being large or bulky. Crowell alleged that the practice of police officers on the day of the parade was to search the bags of anyone who appeared to be demonstrating against the parade, and that there was no indication that police searched any of the non-protestors, such as members of the media. Crowell also demonstrated at trial that a backpack identical to the one he wore on the day of the march, containing the same items, did not have any unusual bulges.

The jury returned a verdict for Officer Dixon, finding that he had not falsely arrested Crowell. Crowell appealed to the Ninth Circuit on the grounds that the district court erred in failing to grant his motion for judgment notwithstanding the verdict. Crowell argued that he had not obstructed a police officer by refusing to consent to a search of his backpack because Officer Dixon did not have probable cause to search the backpack. Officer Dixon did not raise the defense of qualified immunity on appeal.

123. Id. at 836 (internal quotations omitted).
124. Brief for Appellant at 11, Graves (No. 02-35119).
125. Id. at 2–3.
126. Id. at 11.
127. Graves, 339 F.3d at 837. The protestors then filed a motion for judgment notwithstanding the verdict under Federal Rule of Civil Procedure 50 and, in the alternative, moved for a new trial under Federal Rule of Civil Procedure 59. Id. at 837–38. The district court denied both motions. Id.
128. Id. at 833. Graves presents a unique situation in that it involves a motion for judgment notwithstanding the verdict in a case where the court subsequently raised the issue of qualified immunity sua sponte. Id. at 846 n.23. When making a determination of qualified immunity, the court must examine the facts in the light most favorable to the plaintiff. See Saucier v. Katz, 533 U.S. 194, 201 (2001). However, when evaluating a motion for judgment notwithstanding the verdict, the court must look at the facts in the light most favorable to the non-moving party, which in this case is the defendant. Fed. R. Civ. P. 50(a)(1); Graves, 339 F.3d at 846 n.24.
129. Brief for Appellant at 12–13, Graves (No. 02-35119).
130. Graves, 339 F.3d at 845 n.23.
B. The Ninth Circuit Concluded That Officer Dixon Did Not Have
Probable Cause To Search Crowell’s Backpack

In *Graves*, the Ninth Circuit held that even in light of the hostile
circumstances surrounding the Aryan Nations parade, Officer Dixon
did not have adequate individualized suspicion of Crowell to give him
probable cause to search Crowell’s backpack. The *Graves* court
concluded that the fact that Crowell’s backpack appeared to be heavy
and bulging did not give Officer Dixon “substantial” individualized
suspicion about Crowell because many innocent objects can make a
backpack look heavy and bulging. The court acknowledged that
Officer Dixon had good intentions, but noted that “a good motive is not
sufficient to show probable cause.”

C. The Ninth Circuit Held That Officer Dixon Was Entitled to
Qualified Immunity Because the Law Concerning Probable Cause
Determinations Was Not Clearly Established

The Ninth Circuit held that Officer Dixon was not liable for violating
Crowell’s Fourth Amendment civil liberties because he was entitled to
qualified immunity. The court raised the issue of qualified immunity
sua sponte. The *Graves* court reasoned that Officer Dixon was entitled
to qualified immunity because the law had not clearly established how
much weight Officer Dixon could give to the surrounding circumstances
when he was deciding whether he had probable cause to search
Crowell’s backpack. The court concluded that “a reasonable
Fourth Amendment Rights of Protestors

officer . . . could have believed that [the circumstances surrounding the parade] carried enough weight to create probable cause when there was at least some individualized suspicion. 138 The court considered Crowell’s backpack to be sufficiently suspicious to make Officer Dixon’s probable cause determination reasonable under the circumstances. 139

IV. THE NINTH CIRCUIT ERRED WHEN IT HELD THAT THE OFFICER WAS ENTITLED TO QUALIFIED IMMUNITY

In Graves, Officer Dixon determined that there was probable cause to search Crowell’s backpack based on the fact that he was a protestor carrying a heavy backpack. 140 Carrying a heavy backpack at a protest march is ordinary behavior and cannot serve as the basis for individualized suspicion. Because a reasonable police officer would have known that probable cause requires individualized suspicion 141 that cannot be based on ordinary behavior that would implicate many innocent people 142—such as carrying a heavy backpack—Officer Dixon was not entitled to qualified immunity. If the Coeur d’Alene Police Department believed that searching the bags of people attending the Aryan Nations parade was necessary to ensure safety, the Department should have performed searches as part of an organized special needs program.

A. Individualized Suspicion of Crowell Could Not Be Based on His Backpack Because Carrying a Heavy Backpack Is Ordinary Behavior at a Protest March

The Graves court held that Officer Dixon had sufficient reasonable suspicion of Crowell to perform a lawful Terry stop based on the hostile circumstances of the parade and Crowell’s heavy backpack, 143 but that Officer Dixon did not have probable cause to search Crowell’s

138. Id. at 847.
139. Id. at 847–48.
140. Id. at 836.
143. See Graves, 339 F.3d at 843.
The court erred in concluding that Officer Dixon had enough reasonable suspicion to stop Crowell because it relied too heavily on the surrounding circumstances of the parade and placed insufficient emphasis on Officer Dixon’s lack of individualized suspicion of Crowell. Police officers may consider the surrounding circumstances in addition to the individualized suspicion of a particular person when making an assessment of reasonable suspicion or probable cause. However, individualized suspicion is always necessary for police officers to conduct a lawful *Terry* stop or a search under the Fourth Amendment. In *Graves*, the surrounding circumstances of the parade, which included the presence of two organizations known for violence and the report of stolen explosives, were hostile, and thus weighed in favor of the Ninth Circuit’s determination that there was enough reasonable suspicion to stop Crowell. Regardless of these circumstances police officers did not have enough individualized suspicion of Crowell to set him apart from the other parade attendees and allow officers to conduct a lawful *Terry* stop.

Crowell’s backpack cannot be the basis for individualized suspicion because carrying a backpack is ordinary behavior. Officer Dixon based his individualized suspicion of Crowell on the fact that his backpack was heavy and contained cylindrical objects. However, wearing a heavy, bulky backpack is not unusual behavior, especially in the context of a protest march. As the Ninth Circuit noted, wearing a backpack in general is ordinary behavior. At a protest march where food and services are typically not available, people are even more likely to use a bag or backpack to carry necessities such as food, water, and clothing. Testimony from the district court proceedings in *Graves* indicates that many of the protestors at the parade in fact carried some type of bag. For example, Officer Dixon testified that he alone searched the bags of approximately fourteen protestors in addition to Crowell. Other

---

144. *See id.* at 844.
147. *See Graves*, 339 F.3d at 834, 836.
148. *Id.* at 836.
149. *See id.* at 844.
150. *See Brief for Appellant at 11, Graves* (No. 02-35119).
151. *Id.*
Fourth Amendment Rights of Protestors

officers also actively searched protestors’ bags, further indicating that many of the people protesting at the parade carried bags. In addition, while Crowell’s backpack was heavy, there was nothing particularly suspicious about Crowell’s backpack as compared to other bags. At trial, Crowell demonstrated that a backpack identical to the one he was wearing the day of the parade, filled with identical objects, did not have any unusual bulges.

Crowell’s ordinary behavior of carrying a heavy backpack falls far short of the level of individualized suspicion required by the U.S. Supreme Court to establish the requisite degree of reasonable suspicion for a lawful *Terry* stop. For example, in *Reid v. Georgia*, the Court held that officers did not have sufficient individualized suspicion to perform a *Terry* stop of the defendant based on the following factors: (1) traveling from Ft. Lauderdale; (2) arriving early in the morning; (3) not checking luggage; and (4) trying to avoid looking like he was traveling with a companion. The Court concluded that many innocent travelers engage in the behavior described in the first three factors, and thus these factors were not enough for individualized suspicion of the defendant. While noting that the fourth factor may provide some level of individualized suspicion of the defendant, the Court concluded that it was not enough to support reasonable suspicion. As in *Reid*, Officer Dixon’s basis for searching Crowell’s backpack involved ordinary conduct that was likely engaged in by a number of innocent people at the parade—simply carrying a heavy, and perhaps bulky, backpack. Therefore, this factor cannot support individualized suspicion of Crowell. Even if the heaviness or bulkiness of Crowell’s backpack was out of the ordinary, this alone was insufficient to establish individualized suspicion. Carrying a heavy backpack is significantly less suspicious than trying to conceal that one is traveling with a companion—the unusual behavior that the *Reid* Court found insufficient to demonstrate individualized suspicion.

152. *Id.* at 2.
153. *See id.* at 11.
154. *Id.*
155. *Reid v. Georgia*, 448 U.S. 438, 441 (1980); *see also supra* notes 47–54 and accompanying text.
156. *Reid*, 448 U.S. at 441.
157. *Id.*
159. *Reid*, 448 U.S. at 441.
In addition, the level of suspicion that attached to Crowell’s backpack is much less compelling than the evidence the U.S. Supreme Court held to be sufficient to establish individualized suspicion of an alleged drug courier in United States v. Sokolow. In Sokolow, officers performed a Terry stop of the defendant based on several factors that, although not illegal, were highly unusual and would be unlikely to include many innocent individuals. These factors included paying for plane tickets with $2100 in twenty-dollar bills and traveling under an alias. In contrast, Crowell’s behavior of wearing a heavy backpack was entirely ordinary conduct that was very likely to be engaged in by a number of innocent individuals. Carrying a heavy backpack is significantly less suspicious than paying for an expensive airplane ticket with low-denomination bills and traveling under a false name. Thus, Officer Dixon’s individualized suspicion of Crowell was based on ordinary behavior that would have implicated a number of innocent individuals.

The Ninth Circuit reached the correct conclusion when it held that Officer Dixon had insufficient individualized suspicion of Crowell to lawfully search his backpack. A lawful Terry stop requires that police officers establish reasonable suspicion, while a lawful search must be based on the more exacting standard of probable cause. Individualized suspicion of a particular person is a requirement of both reasonable suspicion and probable cause. Officer Dixon did not have enough individualized suspicion of Crowell to establish the reasonable suspicion predicate for a Terry stop. Therefore, he also had insufficient individualized suspicion of Crowell to meet the more stringent predicate of probable cause for a search.

B. Officer Dixon Was Not Entitled to Qualified Immunity Because Crowell’s Backpack Did Not Provide Sufficient Individualized Suspicion To Establish Probable Cause

In Graves, the Ninth Circuit held that Officer Dixon was entitled to qualified immunity for unlawfully arresting Crowell. The court

160. United States v. Sokolow, 490 U.S. 1, 3, 8–9 (1989); see also supra notes 62–64 and accompanying text.
161. Id. at 8–9.
162. See id. at 7; Terry v. Ohio, 392 U.S. 1, 30 (1968).
165. See Ybarra, 444 U.S. at 90–91.
Fourth Amendment Rights of Protestors

concluded that the law was not clearly established regarding how much weight Officer Dixon could give to the surrounding circumstances of the parade when making his determination of whether there was probable cause to search Crowell’s backpack.\textsuperscript{166} The \textit{Graves} court erred when it held that Officer Dixon was entitled to qualified immunity because a reasonable officer would have known that regardless of the surrounding circumstances, individualized suspicion is necessary to make a determination of probable cause.\textsuperscript{167} Furthermore, a reasonable officer would have been aware that, in the context of a \textit{Terry} stop or a search, individualized suspicion cannot be based on ordinary factors that would implicate many innocent people,\textsuperscript{168} such as carrying a heavy backpack.

Officer Dixon was not entitled to qualified immunity for performing a \textit{Terry} stop of Crowell because the law was clearly established that individualized suspicion is a necessary prerequisite for a lawful \textit{Terry} stop and Officer Dixon did not have sufficient grounds for individualized suspicion of Crowell. The U.S. Supreme Court has clearly established that a \textit{Terry} stop must be grounded on individualized suspicion, regardless of the surrounding circumstances.\textsuperscript{169} U.S. Supreme Court and Ninth Circuit precedent have also clearly established that, in the context of a \textit{Terry} stop, individualized suspicion cannot be based on ordinary factors that would implicate a number of innocent individuals.\textsuperscript{170} In \textit{Graves}, Officer Dixon based his individualized suspicion of Crowell on the fact that Crowell carried a heavy, bulky backpack,\textsuperscript{171} which is a factor that would implicate a large number of innocent individuals.\textsuperscript{172} He also relied on the hostile surrounding circumstances of the parade in his assessment of reasonable suspicion.\textsuperscript{173} Because the law was clearly established that police officers must have individualized suspicion of a particular person to establish reasonable suspicion,\textsuperscript{174} and that individualized suspicion cannot be based on

\begin{flushleft}
\begin{enumerate}
\item \textsuperscript{166} See \textit{Graves v. City of Coeur d’Alene}, 339 F.3d 828, 847–48 (9th Cir. 2003).
\item \textsuperscript{167} See \textit{Ybarra}, 444 U.S. at 91.
\item \textsuperscript{168} See \textit{Reid v. Georgia}, 448 U.S. 438, 441 (1980); \textit{United States v. Rodriguez}, 976 F.2d 592, 595–96 (9th Cir. 1992).
\item \textsuperscript{169} See \textit{Brown}, 443 U.S. at 51–53.
\item \textsuperscript{170} See \textit{Reid}, 448 U.S. at 441; \textit{Rodriguez}, 976 F.2d at 595–96.
\item \textsuperscript{171} See \textit{Graves}, 339 F.3d at 836.
\item \textsuperscript{172} See \textit{supra} Part IV.A.
\item \textsuperscript{173} See \textit{Graves}, 339 F.3d at 836–37.
\item \textsuperscript{174} See \textit{Brown}, 443 U.S. at 51–53.
\end{enumerate}
\end{flushleft}
ordinary factors that would implicate many innocent people, a reasonable police officer should have known that Crowell’s ordinary behavior of carrying a backpack was insufficient to establish reasonable suspicion of him. Although the surrounding circumstances of the parade were hostile, and favored a finding of reasonable suspicion, the law was clearly established that surrounding circumstances alone are not enough to establish reasonable suspicion. Thus, a reasonable police officer should have known that even under the circumstances of the parade, there was not enough individualized suspicion of Crowell to subject him to a Terry stop.

In addition, Officer Dixon was not entitled to qualified immunity for unlawfully arresting Crowell because a reasonable officer would have known that Crowell’s ordinary behavior of carrying a backpack did not provide enough individualized suspicion of him to establish probable cause for a search. The law was clearly established that probable cause must be based on individualized suspicion, even when the surrounding circumstances weigh in favor of establishing probable cause. Unlike the Terry stop context, there is no case law holding that, in the context of searches, individualized suspicion cannot be based on ordinary factors that would implicate many innocent people. However, case law with the same facts is not necessary for the law to be clearly established. For example, in Mendoza v. Block, the Ninth Circuit held that the law regarding the use of excessive force was clearly established even though there was little case law dealing specifically with the use of police dogs to locate suspects because a reasonable officer would know that using a “weapon,” such as a police dog, would be unlawful in some circumstances. The Mendoza court concluded that the principle that officers cannot use excessive force was clearly established even when applied to a new situation. Likewise, in Graves, a reasonable officer would know that the principle that individualized suspicion for a Terry stop cannot be based on factors that would include a number of innocent individuals also applies to a similar situation—searches. Thus, because

175. See Reid, 448 U.S. at 441; Rodriguez, 976 F.2d at 595.
176. See Graves, 339 F.3d at 836–37.
177. See Brown, 443 U.S. at 51–53.
180. Mendoza, 27 F.3d at 1362.
181. See id.
Fourth Amendment Rights of Protestors

individualized suspicion is a requirement that is common to, and necessary for, both searches and Terry stops; it would have been apparent to a reasonable police officer from existing case law that individualized suspicion cannot be based on factors that would implicate a large number of innocent individuals in the context of a search or Terry stop. Furthermore, because it was clearly established that the predicate for a search requires a higher level of suspicion than the predicate for a Terry stop, it would be apparent to a reasonable officer that the standard for individualized suspicion could not be lower for a search than for a Terry stop.

C. To Protect Protestors and Participants at the Aryan Nations Parade, the Coeur d’Alene Police Should Have Searched Bags Pursuant to a Special Needs Program

If Coeur d’Alene police officers believed that suspicionless searches were necessary to safeguard against people carrying bombs into the parade, they should have employed a system of special needs searches to detect bombs without violating the civil liberties of protestors. The U.S. Supreme Court has established that suspicionless searches are allowable as long as they are part of a special needs program. To be valid under the Fourth Amendment, these programs must be based on special needs, beyond normal law enforcement goals. Police officers must operate the searches based on predetermined, objective criteria, and searches cannot be grounded in suspicion of a particular individual or class of individuals.

In Graves, the Coeur d’Alene Police Department had a valid goal, beyond normal law enforcement needs, for conducting searches of protestors at the parade—preventing violence. However, the searches conducted did not qualify under the special needs standard because they

182. See Ybarra, 444 U.S. at 91.
188. See Graves v. City of Coeur d’Alene, 339 F.3d 828, 834 (9th Cir. 2003).
were not based on predetermined, objective criteria. In contrast to Sitz, where the U.S. Supreme Court upheld an organized random highway screening program for alcohol intoxication, the officers in Graves did not determine which bags to search based on a previously established plan.

In the future, police officers should establish procedures to perform special needs searches if they believe that suspicionless searches of people’s effects are necessary to maintain public safety at demonstrations. By searching bags using a special needs program based on uniform, predetermined criteria, police officers can enhance safety without violating protestors’ civil liberties.

V. CONCLUSION

In Graves v. City of Coeur d’Alene, the Ninth Circuit held that there was insufficient probable cause for the defendant police officer to search and arrest a protestor at an Aryan Nations march. However, the court invoked the doctrine of qualified immunity to protect the officer from civil liability, finding that the law regarding probable cause as applied to the facts of Graves was not clearly established. U.S. Supreme Court precedent requires that police officers have individualized suspicion of the suspect for a search to be lawful under the Fourth Amendment. Because the police officer in Graves impermissibly searched the protestor without sufficient individualized suspicion, in violation of clearly established law, the police officer was not entitled to qualified immunity. Concerns about national security cannot justify the abandonment of the constitutional right to be free from unreasonable searches. When faced with similar circumstances, other circuit courts of appeals should hold that police officers are not entitled to qualified immunity.

189. See id. at 845 n.22.
190. See Sitz, 496 U.S. at 455.
191. See Graves, 339 F.3d at 845 n.22.
CAPITAL PUNISHMENT, PROPORTIONALITY REVIEW, AND CLAIMS OF FAIRNESS (WITH LESSONS FROM WASHINGTON STATE)

Timothy V. Kaufman-Osborn*

Abstract: This Article explores the adequacy of one of the safeguards adopted by many states to ensure that the death penalty is applied fairly, following the reinstatement of capital punishment in 1976. Relying chiefly on evidence drawn from Washington State, this Article asks whether the practice of comparative proportionality review has ensured that there is now a rational basis for distinguishing between those who are sentenced to die and those who are not. An analysis of the trial judge reports employed by the Washington State Supreme Court in reviewing death sentences, as well as the method used by the court in conducting its reviews over the course of the past two decades, indicates that the death penalty remains arbitrary and capricious in its administration. The failure of comparative proportionality review furnishes yet another reason for concluding that capital punishment cannot be conducted in a way that comports with claims of fairness.

"[C]apital punishment [must] be imposed fairly, and with reasonable consistency, or not at all."1

INTRODUCTION................................................................................ 776

I. COMPARATIVE PROPORTIONALITY REVIEW: ITS JUSTIFICATION AND CONSTITUTIONAL HISTORY....... 784

II. THE LOGIC OF COMPARATIVE PROPORTIONALITY REVIEW.................................................................................... 794
   A. Determining the Universe of Cases for Comparison........ 795
   B. Specifying the Pool of Similar Cases ................................. 798
   C. Identifying a Test of Proportionality .................................. 802

III. COMPARATIVE PROPORTIONALITY REVIEW IN WASHINGTON ........................................................................ 806

IV. THE DEFICIENCIES OF COMPARATIVE PROPORTIONALITY REVIEW IN WASHINGTON ............ 814
   A. Washington’s Trial Judge Reports and Their Defects...... 815
      1. Absent, Unrevised, and Late Reports ........................... 816
      2. Inaccurate and/or Inadequate Information Regarding Defendants and Victims............................................. 822

* Baker Ferguson Professor of Politics and Leadership, Whitman College. For their assistance at various stages of this project, I would like to thank Beth Andrus, Dave Beckley, Alissa Berley, Ashifi Gogo, Mark Larranaga, Linh Ngo, Rebecca Lindemann, Jerry Sheehan, and Nancy Talner.

INTRODUCTION

The United States is now immersed in the most urgent debate about capital punishment in decades. This controversy is most often framed, explicitly or implicitly, through reference to the concept of fairness. The debate encompasses a broad range of questions, including the possible execution of the innocent, the deficient legal representation afforded those charged with capital crimes, the effective restriction of the death penalty to those who are impoverished, the racially discriminatory administration of capital punishment, and the execution of those who were minors at the time they committed their crimes. Each of these
questions raises concerns about whether the death penalty is now (or can ever be) applied in a way that is equitable.

Recently, yet another question regarding the fairness of the death penalty’s administration has received considerable media attention. That question, phrased in the language of proportionality, was raised in dramatic fashion by the plea bargain struck by the so-called “Green River killer,” Gary Ridgway. On November 5, 2003 in a Seattle courtroom, Ridgway confessed to murdering forty-eight young women, making him the most prolific serial killer in U.S. history. In return for providing information to the sheriff’s office about these killings, the prosecuting attorney for King County agreed not to seek the death penalty. Not surprisingly, this decision provoked considerable public furor, exemplified by the response of Mike Carrell, a Republican member of the Judiciary Committee of the Washington State House of Representatives: “If Ridgway does not deserve the death penalty, then what situation and who does deserve it?” If Ridgway, with a body count of four dozen, is not condemned to die, then how can one justify sentencing to death a defendant who has murdered far fewer? How can one explain the Ridgway plea bargain to the ten persons now on death row in Washington, a majority of whom were convicted of murdering a single individual? Some speculate that these questions of fairness will make it more difficult to persuade courts in Washington, and perhaps elsewhere, to sentence those convicted of far less heinous crimes. Indeed, the executive director of the Death Penalty Information Center, Richard Dieter, has suggested that legal challenges predicated on the


5. Id.
6. Id.
Ridgway case may in time attract the attention of the U.S. Supreme Court: “There will be appeals, and there may well be a review about the whole country’s use of the death penalty.”

The roots of the controversy provoked by the Ridgway deal, as well as the broader question of proportionality, can be traced to Furman v. Georgia. Consisting of a brief per curiam opinion, followed by nine separate opinions in a five to four vote, the Furman Court invalidated Georgia and Texas statutes that permitted defendants to be sentenced to death at the unfettered discretion of a judge or jury. The Court held that these laws, and by extension the death penalty statutes of thirty-seven additional states along with various federal statutory provisions, constituted cruel and unusual punishment in violation of the Eighth and Fourteenth Amendments. In his concurring opinion, Justice Stewart maintained that a legal order that allows death sentences to be “wantonly” and “freakishly” imposed is “cruel and unusual in the same way that being struck by lightning is cruel and unusual.” Justice White expressed much the same point, stating that “there is no meaningful basis for distinguishing the few cases in which it is imposed from the many cases in which it is not.” Justice Douglas, endorsing the constitutional conclusion of Justices Stewart and White on somewhat different grounds, explained that the death penalty statutes in question were unconstitutional not because their administration was capricious, but rather because their application to particular categories of persons was all too predictable: “We know that the discretion of judges and juries in imposing the death penalty enables the penalty to be selectively applied, feeding prejudices against the accused if he is poor and despised, and lacking political clout, or if he is a member of a suspect or unpopular minority.” Regardless of the rationale, the Furman Court

9. Johnson, supra note 8, at A8. Even prior to announcement of the Ridgway deal, attorneys for two defendants accused of killing a woman in Snohomish County sought and secured a delay in prosecutors’ decision whether to seek the death penalty against their clients. See Ian Ith, Ridgway Deal Could Change Criteria for Death Penalty, SEATTLE TIMES, Oct. 17, 2003, at A1. The attorneys argued that this decision should turn on the results of the Ridgway negotiations. Id. at A15.

10. 408 U.S. 238 (1972) (per curiam).
11. Id. at 253 (Douglas, J., concurring).
12. Id. at 239–40.
13. Id. at 309–10 (Stewart, J., concurring).
14. Id. at 313 (White, J., concurring).
15. Id. at 255 (Douglas, J., concurring).
Conduct of Comparative Proportionality Review

required states to incorporate procedural safeguards designed to regulate the otherwise standardless deliberations of judges and juries. Absent such checks there can be little reason to believe that capital punishment, the most severe sanction prescribed by law, is reserved for the most blameworthy defendants found guilty of the most heinous offenses.

Today, some three decades after Furman, but just short of the thirty-year anniversary of Gregg v. Georgia, in which the U.S. Supreme Court authorized the states to resume capital punishment, many of the concerns expressed by the Court in 1972 are being voiced once again. Perhaps the most dramatic echo of these worries was heard in Illinois when, on January 31, 2000, Governor George Ryan imposed a moratorium on executions in that state. His decision to do so, he explained, was prompted by questions about the fairness of capital punishment. Specifically, he noted that while twelve inmates had been executed since the death penalty was reinstated in Illinois in 1977, thirteen had been exonerated and freed from death row. Shortly after this announcement, Governor Ryan appointed a commission and instructed it to furnish an explanation for this record of error and to advance recommendations aimed at reforming the administration of capital punishment in the state.

Two years later, the governor’s commission called for a sweeping overhaul of the death penalty in Illinois, although its members conceded that “no system, given human nature and frailties, could ever be devised or constructed that would work perfectly and guarantee absolutely that no innocent person is ever again sentenced to death.”

---

16. Id. at 400–01 (Burger, C.J., dissenting) (“While I would not undertake to make a definitive statement as to the parameters of the Court’s ruling, it is clear that if state legislatures and the Congress wish to maintain the availability of capital punishment, significant statutory changes will have to be made.”).


18. Id. at 207 (plurality opinion).


20. Id.

21. Id.


concern, a study appended to the commission’s report reached two conclusions. First, the race of homicide victims is a statistically significant predictor of who is and is not sentenced to death in Illinois. Second, the frequency of death sentences varies markedly from region to region throughout Illinois. Responding to these findings, in January, 2003, Governor Ryan commuted every death sentence then in effect to prison terms of life or less: “The facts that I have seen in reviewing each and every one of these cases,” he declared, “raised questions not only about the innocence of people on death row, but about the fairness of the death penalty system as a whole. Our capital system is haunted by the demon of error: error in determining guilt and error in determining who among the guilty deserves to die.”

Governor Ryan’s worries about the justice of capital punishment in Illinois have been articulated throughout the United States, and major


25. Id. Contrary to the contention of Justice Douglas in Furman, the study by Pierce and Radelet did not find statistically significant evidence of disparate treatment based on the race of the defendant, holding aggravating factors constant. Id. However, the study did caution that, because its inquiry was limited to defendants convicted of first-degree murder and because earlier stages in the judicial process were not examined, “estimates of arbitrariness and/or discrimination . . . may underrepresent the effects of extra-legal factors.” Id.; see also Raymond Paternoster, An Empirical Analysis of Maryland’s Death Sentencing System with Respect to the Influence of Race and Legal Jurisdiction 33 (2003) (finding that black killers of white victims are nearly three and a half times as likely to be sentenced to die as are blacks who kill blacks in Maryland, in a recent study demonstrating significant racial bias in the capital punishment system), available at http://www.urhome.umd.edu/newsdesk/pdf/finalrep.pdf (last visited June 27, 2004); U.S. Dep’t of Justice, The Federal Death Penalty System: A Statistical Survey (1988-2000) (2000) (conducting a comparable inquiry into racial bias and geographic disparity in the administration of the federal death penalty statute), available at http://www.usdoj.gov/dag/pubdoc/dpsurvey.html (last visited June 27, 2004).

Conduct of Comparative Proportionality Review

death penalty studies have recently been conducted in Arizona, Connecticut, Florida, Illinois, Indiana, Maryland, Nebraska, Nevada, New Jersey, North Carolina, Virginia, and at the federal level.\(^27\) In addition, legislatures in thirty-seven of the thirty-eight states that now provide for the death penalty considered reforms to its administration during their legislative sessions in 2001, with twenty-one of those states adopting at least one such reform.\(^28\) Statements by public figures and in opinion polls also show a growing concern about the death penalty’s fairness. For example, several long-time supporters of the death penalty, including George Will, the Reverend Pat Robertson, and Oliver North have expressed reservations about the death penalty based on concerns about the fairness of its administration.\(^29\) Justice Sandra Day O’Connor, a longtime advocate of the death penalty, delivered two speeches in 2001 in which she noted that “serious questions” have been raised in recent years “about whether the death penalty is being fairly administered in this country.”\(^30\) Attesting to the cumulative effect of the publicity afforded to such concerns, a Gallup poll conducted in 2000 indicated that a bare majority of Americans (51%) believe that the death penalty is applied fairly, while 41% believe that it is not.\(^31\) Moreover, support for the death penalty in national polls has dropped from about 77% to 63% since 1996.\(^32\) In 2000, while 53% of Americans “favored a nationwide suspension of executions until a study is completed on the fairness of


\(^{29}\) See ABC This Week: Roundtable Discussion of the Death Penalty Moratorium (ABC television broadcast, Apr. 9, 2000) (discussing the death penalty, George Will expressed his conviction that innocent persons have been executed, and Pat Robertson expressed his support for a moratorium on executions because of concern about capital punishment’s unfairness to the poor and minorities); see also Robert Reno, Support for Death Penalty Goes Wobbly, DES MOINES REG., June 12, 2000, at 7A (quoting Oliver North: “I think capital punishment’s day is done in this country, I don’t think it’s fairly applied.”).


how the death penalty is used,” only 29% opposed this recommendation.33

Together, these concerns about the death penalty’s fairness pose questions about whether the procedural reforms adopted by the states after 1972 and approved by the U.S. Supreme Court in 1976 have rendered the death penalty any less arbitrary and capricious than it was deemed in Furman. A mounting body of evidence suggests that the safeguards affirmed in Gregg do not in fact ensure the equitable administration of the death penalty demanded by Furman.34 If that is so, then it is hard to escape the conclusion of James Liebman and his associates, authors of a comprehensive national study of error rates in capital cases: “the time has come to fix the death penalty, or end it.”35

The purpose of this Article is to examine the efficacy of one of the less frequently discussed safeguards commended by the Gregg plurality in authorizing the reinstatement of capital punishment.36 Specifically, this Article evaluates the adoption and implementation of comparative proportionality review throughout Washington’s judicial system. Comparative proportionality review is a statutorily mandated procedure intended to combat the two problems that arise when excessive discretion is granted to a jury vested with the authority to impose the death penalty.37 By comparing any given death sentence with the penalties imposed on others convicted of death-eligible crimes, comparative proportionality review is intended to serve two purposes.38 First, it is intended to ensure that there is a rationally defensible basis for distinguishing those sentenced to die from those who are not.39 Second, it is intended to prevent death sentences predicated on constitutionally impermissible factors such as economic status or racial identity, whether

34. Furman v. Georgia, 408 U.S. 238, 256 (1972) (Douglas, J., concurring); see supra note 3 (citing articles examining some of the mounting evidence against the equitable administration of the death penalty).
37. Id. at 188–89 (plurality opinion) (expressing concern about excessive discretion on the part of sentencers).
38. Id. at 204–06 (plurality opinion) (discussing comparative proportionality review as a remedy for prejudice and arbitrariness).
39. Id.
Conduct of Comparative Proportionality Review

of the defendant or the victim. 40
The evidence adduced in this Article is drawn chiefly from Washington, and many of its claims are specific to that state’s comparative proportionality review. However, the flaws identified by this analysis suggest a broader conclusion: the implementation of comparative proportionality review over the last three decades has not provided, and indeed cannot provide, an adequate safeguard against the arbitrary and capricious administration of capital punishment. The predicaments encountered by the Washington State Supreme Court in its conduct of comparative proportionality review are indicative and, in part, constitutive of the defining dilemma that courts have struggled to resolve over the course of the past three decades in attempting to craft a constitutionally coherent doctrine regarding capital punishment. 41
Specifically, the deficiencies of comparative proportionality review in Washington mirror the judiciary’s difficulties in fashioning a principled standpoint that will ensure that the death penalty is administered in a way that meets the standards of fairness in two distinct and arguably incompatible senses of the term. First, administration of the death penalty must be fair in the sense that it must be restricted to, as well as consistently applied to, only the most heinous criminals found guilty of the most heinous crimes. 42 Second, it must be fair in the sense that the judicial processes that generate, review, and affirm death sentences must provide full consideration to the individualized character and circumstances of each capital defendant. 43 Insofar as the practice of comparative proportionality review in Washington founders on the competing claims of these two understandings of fairness, this Article indicates why, of the two options proposed by Liebman and his associates, to fix or to abolish the death penalty, only the latter is compatible with the diverse imperatives of justice.

Parts I and II of this Article offer two contexts for understanding the

40. Id.
42. See Gregg, 428 U.S. at 187 (plurality opinion) (stating that capital punishment is “an extreme sanction, suitable to the most extreme of crimes”).
43. Woodson v. North Carolina, 428 U.S. 280, 304 (1976) (plurality opinion) (noting that a “process that accords no significance to relevant facets of the character and record of the individual offender or the circumstances of the particular offense excludes from consideration in fixing the ultimate punishment of death the possibility of compassionate or mitigating factors stemming from the diverse frailties of humankind”).

783
practice of comparative proportionality review as it is currently conducted in Washington. Part I examines the ethical and legal justifications as well as the constitutional history behind this practice. Part II looks at the logic of comparative proportionality review, i.e., the principal conceptual questions that must be answered in order to translate an abstract statutory commitment into a coherent judicial practice. Part II.A addresses the question of how to define the universe of cases deemed relevant to the conduct of such review. Part II.B addresses the question of how to extract from this universe the smaller pool of cases considered similar to any given case on appeal. Part II.C addresses the question of how to designate the criteria necessary to generate and sustain the conclusion that a death sentence under review is or is not proportionate. Part III then examines the history of comparative proportionality review’s adoption in Washington. Part IV analyzes the principal deficiencies of comparative proportionality review as currently practiced. Specifically, Part IV.A examines deficiencies in the database of cases that the Washington State Supreme Court is statutorily required to consider in conducting comparative proportionality review, and Part IV.B critiques the way in which the court has actually conducted its proportionality reviews. Finally, Part V returns to the broader implications of comparative proportionality review for the current controversy regarding the fairness of capital punishment in contemporary America.

I. COMPARATIVE PROPORTIONALITY REVIEW: ITS JUSTIFICATION AND CONSTITUTIONAL HISTORY

Within the context of capital punishment jurisprudence, comparative proportionality review is a procedure by which a court determines whether a death sentence is consistent with the usual pattern of sentencing decisions in similar cases.

44. See Gregg, 428 U.S. at 204–05 (plurality opinion).

Conduct of Comparative Proportionality Review

Amendment, fairness dictates that like crimes should be punished alike. By extension, it is unfair to mete out harsher punishment to certain members of a class of persons convicted of the same crime. Conversely, it is unfair to impose a less harsh punishment on some convicted of the same crime (leaving aside for now the complications involved in determining what qualifies as the “same” crime).

Determination of consistency in the context of capital punishment necessarily involves assessing the proportionality of a death sentence by comparing it to similar cases in which the death penalty has or could have been imposed. Granted, in particular cases, comparative analysis may not be imperative in order to determine whether a death sentence has been imposed as a result of constitutionally impermissible considerations such as race. Such comparative analysis, however, is a crucial means of determining whether a systematic pattern of arbitrary (in the sense of discriminatory) sentencing exists in any given jurisdiction and hence, whether fairness (in the sense of equal treatment) has been achieved.

A general commitment to proportionality as a vital ingredient of fairness is expressed in the text of the U.S. Constitution. The Eighth Amendment prohibits “excessive” bail and fines, and proscribes “unusual” punishments. Obviously, neither what is “excessive,” nor what is “unusual,” can be determined absent consideration of some norm from which such punishments are said to depart. What is less apparent is whether the U.S. Constitution specifically requires comparative proportionality review or, as some have argued, only “inherent” proportionality review. The latter requires that any given punishment be commensurate to the offense in question.

46. U.S. Const. amend. XIV.
47. Radin, supra note 45, at 1166.
48. Id.
50. U.S. Const. amend. VIII.
punishment may be deemed unconstitutional because it is excessive per se, without reference to sentences imposed on others for the same crime.54 For example, in Coker v. Georgia,55 a plurality of the U.S. Supreme Court concluded that the death penalty is always disproportionate to the crime of rape of an adult.56 Justifying this conclusion, the Court first asked whether the death penalty, when imposed for this crime, advances either of the constitutionally permissible purposes of capital punishment: retribution or deterrence.57 Unless the death penalty significantly contributes to one or both of these goals, as the Court later elaborated in Enmund v. Florida,58 “it ‘is nothing more than the purposeless and needless imposition of pain and suffering,’ and hence an unconstitutional punishment.”59 Second, the Court asked whether the punishment imposed, “is grossly out of proportion to the severity of the crime.”60 Because rape, unlike murder, does not involve the taking of life, the Coker plurality concluded that death is an excessive and hence unconstitutional penalty for this category of offense.61

Unlike the Court in Coker, the Furman Court’s principal concern was not that of the inherent disproportionality of death as a punishment for a particular category of crimes, but rather the comparative inconsistency between sentences imposed on different persons found guilty of similar crimes.62 Because of the unbridled discretion vested in judges and juries, existing capital punishment statutes generated a pattern of death

54. See Trop, 356 U.S. at 100–01; Weems, 217 U.S. at 382.
55. 433 U.S. 584 (1977). There are two principal antecedents to Coker, each of which rests on the concept of inherent disproportionality. See Trop, 356 U.S. at 100–01 (holding that, according to “the evolving standards of decency that mark the progress of a maturing society,” deprivation of citizenship is an excessive punishment for a conviction imposed by a court martial for the offense of wartime desertion); Weems, 217 U.S. at 382 (holding that fifteen years of incarceration with hard labor is an excessive punishment for the crime of falsifying a public document).
56. Coker, 433 U.S. at 592 (plurality opinion).
57. Id. (citing Gregg v. Georgia, 428 U.S. 153, 183 (1976) (plurality opinion)).
59. Id. at 798 (citation omitted).
60. Coker, 433 U.S. at 592 (plurality opinion). Arguably, the Coker plurality advanced a second rationale for its conclusion as well. In 1977, when Coker was decided, only three states authorized the death penalty for rape. Id. at 595–96 (plurality opinion). In two of these states, the victim had to be a child in order to warrant a death sentence. Id. On this basis, the Court concluded that the penalty of death for rape was no longer consistent with public opinion, as reflected in state statutes, and, as such, was no longer congruent with contemporary norms of appropriate punishment. Id.
61. Id. at 598 (plurality opinion).
sentencing, which, according to Justice Brennan, “smacks of little more than a lottery system.”63 In part, that “lottery” is problematic because, as Justice Douglas suggested, one can explain why some are sentenced to death while others are not only by citing forms of discrimination that offend the principles of equality under the law.64 However, it is also problematic because it may generate too many as well as too few death sentences.65 On the one hand, Furman implied a concern about the problem of overinclusion. If death penalty statutes fail to provide juries adequate guidance in determining what counts as the most death-worthy offenses, an offender may be sentenced to death for a crime that does not so qualify.66 On the other hand, Furman also implied a concern about the problem of underinclusion. If the death penalty is imposed in only a small number of the cases for which it is legally available, it may prove difficult, if not impossible, to articulate a coherent basis for differentiating between the few who are sentenced to death and the many who receive life sentences.67 As Justice Brennan summarized, “[w]hen the punishment of death is inflicted in a trivial number of the cases in which it is legally available “the conclusion is virtually inescapable that it is being inflicted arbitrarily.”68

The immediate effect of Furman was to prompt state legislatures to revise their death penalty statutes in an effort to address the Court’s concerns about excessive discretion.69 The Court first evaluated the constitutionality of these efforts in Woodson v. North Carolina.70 In that case, the Court struck down a North Carolina statute that mandated the death penalty for everyone convicted of first-degree or felony murder.71 While the statute was intended to eliminate inconsistent sentencing patterns, Justice Stewart concluded that such a mandatory sentencing
provision does not allow for “particularized consideration of relevant aspects of the character and record of each convicted defendant before the imposition upon him of a sentence of death.”72 Predicated on the belief that the penalty of death is qualitatively different from a sentence of imprisonment, however long, the Woodson Court insisted that “individualiz[ed] sentencing” is a “constitutionally indispensable part of the process of inflicting the penalty of death.”73

In contrast, in Gregg v. Georgia74 the Court found Georgia’s revised statute to be constitutional.75 The Georgia legislature sought to meet Furman’s concerns by adopting three reforms.76 First, it bifurcated capital proceedings into an initial phase to determine guilt or innocence and a subsequent phase to determine the appropriate sentence to be imposed upon those found guilty.77 Second, in order to guide jury deliberations, it required that juries consider and, if applicable, weigh statutorily prescribed aggravating factors against mitigating factors before determining the appropriateness of a death sentence.78 Third, and of principal interest here, it mandated appellate review of all death sentences, including the requirement that the Georgia State Supreme Court conduct comparative proportionality analyses.79

More precisely, the Georgia law required its highest court to determine whether any given death sentence “is excessive or disproportionate to the penalty imposed in similar cases, considering both the crime and the defendant.”80 Should the court determine that

72.  Id. at 303 (plurality opinion). At the same time, a plurality of the Court found unconstitutional a Louisiana statute, which, like its North Carolina counterpart, also required the imposition of death sentences for certain crimes. Roberts v. Louisiana, 428 U.S. 325, 329 (1976) (plurality opinion). In addition to arguing that such a procedure fails to allow particularized consideration of any given defendant’s character and record, the Woodson and Roberts pluralities concluded that this requirement offends contemporary standards of decency and perpetuates the concerns about arbitrariness identified in Furman, specifically by encouraging the practice of jury nullification. Id. at 347 (plurality opinion); Woodson, 428 U.S. at 293 (plurality opinion).
73.  Woodson, 428 U.S. at 304 (plurality opinion).
75.  Id. at 207 (plurality opinion).
76.  Id. at 153–54 (plurality opinion).
80.  Id.
Conduct of Comparative Proportionality Review

Georgia juries do not generally impose the death penalty in factually similar cases, even if otherwise properly imposed, that sentence was to be set aside and the case remanded for resentencing. Commending this requirement, the Gregg plurality predicted:

The provision for appellate review in the Georgia capital-sentencing system serves as a check against the random or arbitrary imposition of the death penalty. In particular, the proportionality review substantially eliminates the possibility that a person will be sentenced to die by the action of an aberrant jury. If a time comes when juries generally do not impose the death sentence in a certain kind of murder case, the appellate review procedures assure that no defendant convicted under such circumstances will suffer a sentence of death.

In addition, the plurality commended the statutory provision that required trial judges to complete and submit a standard questionnaire to the state’s highest court within ten days of sentencing. That questionnaire required the trial judge to determine the influence, if any, of “passion, prejudice, or any other arbitrary factor” in the deliberations of the sentencing authority and, in addition, to assess whether the evidence adduced at trial sustains the finding of at least one aggravating circumstance, which is a statutory prerequisite for the death sentence’s imposition. Finally, the Georgia statute required the state’s highest court to include express “reference to those similar cases which it took into consideration” in rendering its judgment regarding proportionality, thereby ensuring that its rationale would be expressly

81. See id.
82. Gregg v. Georgia, 428 U.S. 153, 206 (1976) (plurality opinion). Additionally, Justice White, joined by Chief Justice Burger and Justice Rehnquist, predicted that “if the Georgia Supreme Court properly performs the task assigned to it under the Georgia statutes, death sentences imposed for discriminatory reasons or wantonly or freakishly will be set aside.” Id. at 224 (White, J., concurring).
83. Id. at 167–68 (plurality opinion).
85. Id. (codified at GA. CODE ANN. § 17-10-35(c)(2) (1996)).
86. Id. (codified at GA. CODE ANN. § 17-10-35(e) (1996)). The Court’s commitment to the importance of using comparative proportionality review was effectively reaffirmed in Zant v. Stephens, 462 U.S. 862, 876 (1983). In Zant, the Georgia death penalty statute was once again upheld, in part because of its comparative proportionality review provision. Id. at 876–77. At issue was whether invalidation of one of the three aggravating factors supporting the death sentence invalidated the sentence imposed. Id. at 864. The Court concluded that the narrowing function performed by the two valid factors adequately distinguished the case under review from other
articulated.

In the wake of *Gregg*, virtually all states whose death penalty statutes were invalidated by *Furman* re-crafted their laws, with many adopting the Georgia statute as their model. Indeed, twenty-six states enacted reformed statutes that included comparative proportionality review provisions similar or, in many cases, identical to those adopted by the Georgia legislature. Three additional states adopted comparative proportionality review as a result of state court decisions. However, while *Gregg* established that Georgia’s death penalty statute was sufficient to meet the concerns articulated in *Furman*, it did not resolve the question of whether any or all of the specific procedural safeguards adopted by that state were necessary in order to render any given capital statute consonant with the U.S. Constitution and, more particularly, the Eighth Amendment.

The Court answered that question in *Pulley v. Harris*, in 1984, holding that comparative proportionality review of death sentences is not constitutionally required. Writing for the Court, Justice White acknowledged that the *Gregg* plurality had “made much” of Georgia’s adoption of comparative proportionality review. However, he insisted, the Court had not declared “that comparative review was so critical that without it the Georgia statute would not have passed constitutional muster.” Moreover, Justice White noted that when the Court decided *Gregg*, it simultaneously approved a revised Texas death penalty statute

Georgia murder cases in which the death penalty had not been imposed. *Id.* at 879. Moreover, the Court upheld the sentence on the grounds that the Georgia State Supreme Court had reviewed the sentence to determine whether it was disproportionate. *Id.*

87. See Latzer, *supra* note 51, at 1168 n.29 (listing the states that adopted comparative proportionality review by statutory reform during the decade following *Furman*).

88. *Id.*


92. *Id.* at 43–44.

93. *Id.* at 45.

94. *Id.* Justices Brennan and Marshall dissented on the grounds that comparative proportionality review has been shown to “eliminate some, if only a small part, of the irrationality that currently surrounds the imposition of the death penalty,” and thus should be required by the U.S. Constitution. *Id.* at 60–61 (Brennan, J., dissenting).
Conduct of Comparative Proportionality Review

that did not include a provision for comparative proportionality review. While the California statute at issue in Pulley did not provide for comparative proportionality review, it did endeavor to limit the discretion of juries. Specifically, the statute “minimize[d] the risk of wholly arbitrary and capricious action” by bifurcating capital trials into separate guilt and sentencing phases, as well as by requiring consideration of aggravating and mitigating factors. Justice White acknowledged that a statute of the sort approved in Pulley “may occasionally produce aberrational outcomes,” but insisted that “such inconsistencies are a far cry from the major systemic defects identified in Furman.” Therefore, while every valid death penalty statute must incorporate some “means to promote the evenhanded, rational, and consistent imposition of death sentences,” those means need not include comparative proportionality review.

Shortly after Pulley was decided, nine states repealed their statutory comparative proportionality review requirements; and several others that had been required to adopt comparative proportionality review by state supreme court mandate abandoned the practice as well. On this basis,

96. Pulley, 465 U.S. at 45.
97. Id. at 53.
98. Id. at 54.
99. Id. at 49 (citation omitted).
100. The Court’s retreat from its endorsement of comparative proportionality review in Gregg was further cemented by its ruling in McCleskey v. Kemp, 481 U.S. 279 (1987). In McCleskey, a black man convicted of murdering a white police officer sought habeas corpus relief in federal court, alleging that the Georgia capital sentencing process was administered in a racially discriminatory manner and, as such, violated the Eighth and Fourteenth Amendments. Id. at 286. In support of this claim, McCleskey offered a statistical study on the imposition of the death sentence in Georgia that demonstrated significant disparities based, most significantly, on the race of the murder victim. Id. at 286–87. The U.S. Supreme Court rejected McCleskey’s claim, holding that in order to prevail he was required to prove that the judge or jury in his particular case had acted with discriminatory purpose. Id. at 297–98. The net effect of this decision was to undercut one of the two central rationales for comparative proportionality review: that of preventing the sort of prejudicial decision-making that can only be determined by comparing any given sentence with the overall pattern of death sentencing in the relevant jurisdiction.
Leigh Bienen concluded that, following *Pulley*, “the majority of state high courts reduced proportionality review to a perfunctory exercise.” Moreover, of the state appellate decisions in capital cases rendered between 1975 and April 1996, only fifty-five death sentences were vacated on the ground of disproportionality, while 1376 death sentences were affirmed. Despite this record, twenty of the thirty-eight states that provide for capital punishment continue to require comparative proportionality review by statute. In addition, the State Supreme Courts of Florida and Arkansas continue to incorporate comparative proportionality review into their death sentence reviews, although not required to do so by statute.

Neither the infrequent and often superficial conduct of such review, nor the lack of an Eighth Amendment mandate for such review, answers the question of whether the determination of comparative proportionality is crucial in order to fashion a system of capital punishment that is consonant with the claims of fairness. That question acquires additional urgency in light of the contemporary controversy over the death penalty. One of the principal recommendations of the commission on capital punishment convened by Governor Ryan is that “the Illinois Supreme Court should consider on direct appeal . . . whether the sentence of death was excessive or disproportionate to the penalty imposed in similar cases.”


105. See Wallace & Sorenson, *Nationwide Examination*, supra note 103, at 16 (providing a table indicating the status of comparative proportionality review requirements in all death penalty states as of 1997).
Conduct of Comparative Proportionality Review

cases.”106 Echoing the statutory provisions commended in Gregg, the commission also urged that information “be collected at the trial level with respect to prosecutions of first-degree murder cases, by trial judges, which would detail information that could prove valuable in assessing whether the death penalty is, in fact, being fairly applied.”107 Among other recommendations, the national study of error rates in capital cases conducted by Liebman and his colleagues concluded by exhorting each jurisdiction to consider “using comparative review of murder sentences to identify what counts as ‘the worst of the worst’ in the state, and overturning outlying death verdicts.”108 Finally, in 2001, The Constitution Project, a bipartisan, nonprofit organization, recommended:

Every state should adopt procedures for ensuring that death sentences are meted out in a proportionate manner to make sure that the death penalty is being administered in a rational, non-arbitrary, and even-handed fashion, to provide a check on broad prosecutorial discretion, and to prevent discrimination from playing a role in the capital decision-making process.109

In sum, we are now in the midst of a national debate about capital punishment fueled by renewed articulations of the concerns, initially advanced in Furman, about the arbitrary and/or discriminatory imposition of the death penalty. In response to these concerns, comparative proportionality review is once again being offered as a possible remedy for that worry. Especially in light of the Ridgway plea bargain, challenges in future murder cases will almost certainly be mounted on the grounds of disproportionality. For all of these reasons, it behooves us to ask whether in principle or in practice such review can be conducted in a way that effectively responds to the concerns about fairness that first prompted its adoption.

106. GOVERNOR’S COMMISSION, supra note 23, at Recommendation 70.
107. Id. at Recommendation 84.
109. See THE CONSTITUTION PROJECT, MANDATORY JUSTICE: EIGHTEEN REFORMS TO THE DEATH PENALTY 27 (2001); see also RACHEL KING, BROKEN JUSTICE: THE DEATH PENALTY IN VIRGINIA 55 (Nov. 2003) (recommending, among other reforms, that the “Virginia Supreme Court should keep a database of all murder convictions—those where the death penalty was imposed and those where it was not—to use for proportionality reviews,” in a study by the American Civil Liberties Union of Virginia and the ACLU Capital Punishment Project endorsed by groups as diverse as the Rutherford Institute and Amnesty International USA), available at http://www.acluva.org (last visited July 29, 2004).
II. THE LOGIC OF COMPARATIVE PROPORTIONALITY REVIEW

There is no single model of comparative proportionality review to which all state appellate courts adhere, and, as indicated below, no model is without problems. That said, the logic of this task requires a jurisdiction to perform three basic steps in order to implement comparative proportionality review. First, a court must select the universe of cases to be considered when such reviews are conducted. Second, a court must choose the pool of cases deemed “similar” to a specific case on appeal. Third, a court must decide whether a specific case is proportionate when measured against the pool of similar cases. Because any judgment about the comparative proportionality of a death sentence is relative in the sense that it is determined through reference to whatever other cases are included for the purposes of comparison, resolution of the first two steps will go a long way toward dictating the results of the third. This also means that any problems in a jurisdiction’s resolution of each of the earlier steps will taint those that follow.

111. Id. at 11–12.
112. Id. at 10–11.
Conduct of Comparative Proportionality Review

A. Determining the Universe of Cases for Comparison

Most state statutes provide little or no guidance about how to determine the universe of cases for comparison. Appropriating the language adopted in Georgia, most mandate merely that appellate courts decide whether a particular death sentence is “excessive or disproportionate to the penalty imposed in similar cases, considering both the crime and the defendant.” Accordingly, appellate courts must independently determine the universe of cases relevant to such comparisons. In this task, a court has a number of options. For the sake of illustrating the range of possibilities, as well as what is at stake in the choice of one as opposed to another, this section will outline the most restrictive as well as the most expansive methods of determining the universe of cases for comparison.

Most restrictively, a court can limit its comparative review to only those cases that resulted in death sentences upheld on appeal. Or, somewhat more broadly, it can limit review to all cases advancing to penalty-phase trials, regardless of whether those proceedings did or did not result in death sentences. This method is intuitively appealing...
because, on its face, it seems reasonable to determine the proportionality of any given death sentence by comparing it to other cases deemed sufficiently heinous to warrant the same punishment. That said, this method suffers from an obvious deficiency. To examine only those cases that culminated in death sentences, or those in which this sentence was considered by a jury during the penalty phase, is to deprive an appellate court of any means of calculating the relative frequency with which this penalty is imposed in the larger class of first-degree murders. Doing so defeats the very purpose of comparative proportionality review because an appellate court is thereby deprived of any basis for determining how many persons convicted of first-degree murder were spared the death penalty. The New Jersey State Supreme Court, in \textit{State v. Marshall},\textsuperscript{118} offered the following hypothetical to illustrate this point:

On the assumption that 100 robbery-felony-murder cases are prosecuted as capital crimes, all defendants are convicted and one defendant is sentenced to death, a comparison of the death-sentenced defendant’s punishment with the punishment imposed only on other death-sentenced defendants would exclude from the proportionality-review process the ninety-nine robbery-felony-murder defendants that juries did not sentence to death. Indisputably, the determination whether that single death sentence is disproportionate can be made only by comparing it with the life sentences imposed on the ninety-nine defendants convicted of the same crime.\textsuperscript{119}

Absent such comparison, this method constructs a universe of cases that tilts comparative proportionality review in favor of upholding any given death sentence.\textsuperscript{120} Equally important, it fails to answer the question posed by Justice White in \textit{Furman}: How is one to meaningfully distinguish the few cases in which death sentences are imposed from the many that result in sentences of life imprisonment?\textsuperscript{121}

At the opposite end of the spectrum, in determining the relevant universe of cases, a court may elect to review all cases in which the facts was properly before the sentencing authority for determination”) (citation omitted), \textit{cert. denied}, 519 U.S. 1013 (1996).
\textsuperscript{118} 613 A.2d 1059 (N.J. 1992).
\textsuperscript{119} \textit{Id.} at 1071.
\textsuperscript{120} Sprenger, \textit{supra} note 115, at 738–40 (concluding that states that exclude life sentence cases from the universe of comparative cases engage in the least effective approach to identify disproportionate sentences).
\textsuperscript{121} \textit{See Furman v. Georgia}, 408 U.S. 238, 313 (1972) (White, J., concurring).
Conduct of Comparative Proportionality Review

provide the legal basis for a possible capital prosecution, regardless of whether such prosecution actually ensues.122 The National Center for State Courts recommended such a position in 1984, when it proposed that “the pool of cases for a proportionality review system should contain, as a minimum, all cases in which the indictment included a death-eligible charge, and a homicide conviction was obtained.”123 This method overcomes the principal deficiency of the first because, by not limiting its universe to those crimes that culminate in death sentences, it provides a more inclusive indication of the diverse ways that the law disposes of first-degree murders. Additionally, this method seeks to remedy not merely the problems occasioned by the discretionary authority of juries, but also by the discretion exercised by prosecutors when they choose whether or not to seek the death penalty in any given case.124 Building on the hypothetical cited above, the New Jersey State Supreme Court explained:

Were we to assume that the remaining ninety-nine defendants were prosecuted and convicted of non-capital murder because of prosecutorial decisions not to seek the death penalty, the disproportionality of the single defendant’s death sentence would arise not because of a disproportionate jury determination but because the prosecutorial decision to seek the death penalty was unique. That type of disproportionate death sentence could not be identified by a proportionality-review process that was limited to [capital cases tried to] a penalty phase; it could be

122. New York, for example, has defined a broad universe encompassing some homicide cases that were not capitally prosecuted. See N.Y. JUD. LAW § 211-a (McKinney 2003); N.Y. CT. R. § 510.18 (authorizing collection of case data for every criminal action in which defendant is indicted for first-degree murder).

123. THE NATIONAL CENTER FOR STATE COURTS, USER MANUAL FOR PROTOTYPE PROPORTIONALITY REVIEW SYSTEMS A-8 (1984) [hereinafter USER MANUAL]. The report noted:

In most jurisdictions, this guideline will mean all cases in which the defendant was charged with first degree murder and convicted of first- or second-degree murder or manslaughter. It is to include convictions resulting from a plea of guilty as well as those following a trial, and life sentences resulting from the absence of any aggravating circumstances as well as those stemming from a jury’s apparent determination to exercise mercy after finding a defendant legally eligible for capital punishment.

Id.

124. See Marshall, 613 A.2d at 1070–73 (arguing on behalf of a broad-based universe of cases and explaining that the exclusion of non-capital murder cases, for example, could lead to disproportionality due to prosecutorial discretion); BALDUS, EQUAL JUSTICE, supra note 113, at 211–12 (arguing that a restrictive specification of the universe of similar cases renders prosecutorial discretion immune to judicial oversight).
identified, however, by a universe that included clearly death-eligible homicides that were not prosecuted as capital cases.  

This expansive method of determining the scope of the universe also suffers from serious flaws. For example, a prosecutor’s decision to seek the death penalty may not be the best way to determine which cases should be included in the universe of cases for comparison because, often, for a number of reasons, prosecutors will elect not to seek a death sentence even when the facts of a case render this option available. Courts are therefore put in the position of engaging in independent discovery of the cases they deem death-eligible. As such, this method imposes an exceedingly expensive and time-consuming burden on the judiciary. In addition, because this method requires courts to render this determination absent the full range of information available to prosecutors, it introduces a highly speculative element into a given universe’s composition. Finally, because this method requires that courts second-guess the judgments of prosecutors, it arguably poses constitutional problems regarding the proper separation of powers between the judicial and executive officers of state government and, more particularly, the measure of discretion justifiably left to the latter.

B. Specifying the Pool of Similar Cases

Once courts determine how to specify the relevant universe of cases, they must then establish the criteria that enable determination of the smaller pool of cases considered “similar” to that currently on appeal. As with the question of determining the universe of cases, this section briefly discusses several ways of defining a pool of similar cases.

125. Marshall, 613 A.2d at 1071 (emphasis in original).
126. See generally Latzer, supra note 51 (criticizing this method when examining the practice of the New Jersey Supreme Court).
127. See Latzer, supra note 51, at 1204.
128. Of course, a court can simply duck this second issue by presupposing that all cases in whatever universe it has specified are by definition “similar.” However, that once again defeats the very purpose of comparative proportionality review and thus does nothing to remedy the concerns about the inconsistent application of capital punishment. See infra Part IV.B.2 (discussing the Washington State Supreme Court’s post-1995 conflation of the categories of universe and pool).
129. See USER MANUAL, supra note 123, at A-4 to A-6 (stating the recommendations of the National Center for State Courts on this question). See generally Baldus, supra note 113 (analyzing the effectiveness of different methods of conducting comparative proportionality review); Bienen, supra note 102 (criticizing the conduct of comparative proportionality review on the ground that it provides only an illusion of fairness); Rhonda Hartman, Critiquing Pennsylvania’s Comparative
Conduct of Comparative Proportionality Review

These differ from one another not so much in kind, but, rather, in degree of abstraction from the concrete facts of the case being appealed. Each suffers from significant practical and conceptual difficulties.\(^{130}\)

The most obvious, and perhaps the most intuitive, way to select a pool of similar cases is to identify cases whose facts are the same as the case under review. Literally construed, this method is impossible because no two cases are ever identical. Moreover, if the notion of similarity is construed in strictly factual terms, then a small number of concrete differences will suffice to render any two cases distinguishable, which, in turn, will render a court unable to fulfill its statutory mandate.\(^{131}\) This difficulty may be answered by identifying more general fact patterns that are common to several cases.\(^{132}\) A court, for example, might compare cases that involved invasion of a victim’s private residence, that were

---

\(^{130}\) As with different ways of specifying the universe of cases, these are not the only conceivable methods by which a court might determine “similarity.” Moreover, it is worth noting that many courts have issued proportionality determinations absent any statement of the method employed to determine similarity. See, e.g., DeYoung v. State, 493 S.E.2d 157, 168 (Ga. 1997) (referring without discussion to appendix listing similar cases where the death penalty was upheld), cert. denied, 523 U.S. 1141 (1998); Davis v. State, 660 So. 2d 1228, 1261–62 (Miss. 1995) (referring without discussion to an appendix with list of capital cases the court previously affirmed), cert. denied, 517 U.S. 1192 (1996); Commonwealth v. Uderra, 766 A.2d 334, 342 (Pa. 1998) (making passing reference to statistical data without mention of similar cases).

\(^{131}\) See, e.g., State v. Correll, 715 P.2d 721, 737 (Ariz. 1986) (finding no cases similar on every factual point); Collins v. State, 548 S.W.2d 106, 122 (Ark. 1977) (determining no similar cases because this was the first case reviewed under the new statute); State v. Plath, 313 S.E.2d 619, 630 (S.C. 1984) (“Lacking precisely identical cases with which to compare these verdicts, we are convinced that the sentence of death is neither excessive nor disproportionate in light of this crime and these defendants.”).

\(^{132}\) See, e.g., Cabello v. State, 471 So. 2d 332, 350–51 (Miss. 1985) (concluding that all robbery/murder cases were similar); State v. Lawson, 314 S.E.2d 493, 503 (N.C. 1984) (requiring “roughly similar” fact patterns); Boggs v. Commonwealth, 331 S.E.2d 407, 422 (Va. 1985) (concluding that all murder for hire cases were similar).
unprovoked, and that were preceded by the infliction of severe physical abuse. However, this approach in and of itself does not answer the question of which features are to be deemed sufficiently salient to merit incorporation into a given fact pattern. For example, for the purposes of identifying a pool of cases similar to one on review, is it relevant if all of the victims were young African-American men or if all of the perpetrators used a specific sort of weapon in committing their crimes? Absent specification of the criteria to distinguish between significant and insignificant similarities, such questions cannot be answered.

These problems are compounded when a court takes seriously the standard statutory mandate to consider not just the nature of the crime, but also the character of the defendant. Assessment of the latter typically involves consideration of mitigating factors that may or may not have been introduced during the sentencing phase of a trial. However, any attempt to render defendants comparable on this basis is likely to defy categorization, because defendants are constitutionally permitted to introduce all mitigating factors they deem relevant, including those that are not expressly provided for by statute. Mitigating factors, for example, can include the death of a parent while a defendant was a child, recent loss of a job, or prior service to the community. Should a court elect to ignore mitigating factors because it determines that these difficulties in comparing facts are insurmountable, it will fail to take into account an indispensable measure of a defendant’s culpability.

To overcome these difficulties in identifying a pool of similar cases, courts often interpret the concept of similarity more broadly.


134. See Lockett v. Ohio, 438 U.S. 586, 605 (1978) (holding that any aspect of a defendant’s character or record may be introduced during the sentencing phase of a capital trial); Rhonda G. Hartman, supra note 113, at 893–99 (discussing the difficulties involved in comparing cases on the basis of mitigating factors).

135. See Baldus, Georgia Experience, supra note 133, at 671–72 (discussing the difference between a fact-specific approach to identifying similar cases as opposed to an approach that seeks to render a judgment regarding overall culpability); Baldus, Georgia Experience, supra note 133, at 675–78 (discussing the Georgia State Supreme Court’s struggle with this question); BALDUS, EQUAL JUSTICE, supra note 113, at 285 (discussing the difficulties inherent in fact-specific definitions of similar cases as well as the need for more general measures of defendant culpability); Liebman, supra note 113, at 1438 (discussing the difficulties inherent in fact-specific as opposed to broader methods of defining similar cases); Liebman, supra note 113, at 1442–58 (discussing the Georgia State Supreme Court’s varying definitions of similar cases in terms of “the same crime,” “similar crimes,” and “similar defendants”); Wallace & Sorenson, Nationwide Examination, supra note 103, at 19–20.
Conduct of Comparative Proportionality Review

practice, courts most often accomplish this by identifying the relevant features of a given case with whatever statutorily defined aggravating factors were found applicable at trial. This might include the fact that the offense in question was committed during the course of an armed robbery, that the victim was a police officer, that the offender was previously convicted of one or more homicides, etc. Alternatively, the pool of similar cases is sometimes established not by identifying those that share the same or similar aggravating factors, but, rather, those in which the number of aggravating factors is the same or nearly so. Because each of these approaches abstracts from the concrete circumstances of any given case, all are capable of generating the conclusion that cases differing markedly in strictly factual terms are nonetheless comparable in terms of their overall type or level of aggravation. That determination in turn is taken as an indicator of the degree of heinousness of any given murder and thus the measure of culpability to be ascribed to its perpetrator.

These approaches, however, are insufficient to differentiate aggravated first-degree murder cases in which death sentences are imposed from those in which this sentence is not imposed. The vast majority of cases in which aggravating factors are found applicable do not result in death sentences, no matter what their number, even when those factors are identical to those of cases in which death sentences were imposed. In addition, an approach that defines similarity in terms

136. See Baldus, Georgia Experience, supra note 133, at 670–72; Wallace & Sorenson, Nationwide Examination, supra note 103, at 20; BALDUS, EQUAL JUSTICE, supra note 113, at 201–02.

137. See Baldus, Georgia Experience, supra note 133, at 670–72; Wallace & Sorenson, Nationwide Examination, supra note 103, at 20; BALDUS, EQUAL JUSTICE, supra note 113, 201.

138. See Baldus, Georgia Experience, supra note 133, at 670–72; Wallace & Sorenson, Nationwide Examination, supra note 103, 20; BALDUS, EQUAL JUSTICE, supra note 113, 201–02.

139. See, e.g., State v. Croggins, 716 P.2d 1152, 1158–61 (Idaho 1985) (vacating death sentence imposed on one of two co-defendants on the ground of lesser culpability); State v. Windsor, 716 P.2d 1182, 1193–96 (Idaho 1985) (using culpability measure when comparing intracase sentences); State v. Williams, 287 N.W.2d 18, 29 (Neb. 1979) (employing aggravating factors to determine comparative culpability of different defendants); State v. Gaskin, 326 S.E.2d 132, 147 (S.C. 1985) (“The facts are not the same in any two cases and, accordingly, our review of the facts relate largely to degrees of culpability of the defendants and the viciousness of the killing.”); State v. Carter, 714 S.W.2d 241, 251 (Tenn. 1986) (using an assessment of culpability to distinguish accomplice’s life sentence); Watkins v. Commonwealth, 331 S.E.2d 422, 440 (Va. 1985) (comparing defendant’s future dangerousness and viliness of the crime with previous defendants).

140. See James S. Liebman, The Overproduction of Death, 100 COLUM. L. REV. 2030, 2053–57 (2000) (stating that on average only three hundred of the approximately 21,000 homicides committed in the United States each year result in death sentences, and that of the 6700 persons
of statutorily defined factors is likely to disregard those which, although not specified by law, a jury may find relevant in determining the sentence. Hence, a focus on aggravating factors alone will not take into account a defendant’s willingness to cooperate with authorities or a defendant’s remorsefulness, while an approach that seeks to overcome this difficulty by discovering what factors actually influenced sentencing decisions is beset by its own difficulties. For example, while one jury may conclude that prior drug use renders a crime less excusable, another may deem it a mitigating factor. By the same token, the existence of an intimate relationship between offender and victim may lead one jury to conclude that a murder is more inexplicable and so more heinous, while another jury may conclude that such a relationship renders a murder more understandable and hence its perpetrator more worthy of mercy.

In sum, approaches to the identification of a pool of similar cases that rely on strictly factual similarity founder due to their concrete specificity, whereas approaches that move away from the imperatives of strictly factual congruence stumble over the pitfalls of abstraction. The latter enables a court to assemble a group of cases for the purpose of engaging in comparative proportionality review. However, in doing so, it risks glossing over the facts that distinguish a case on review from other members of the group to which it belongs, thereby jeopardizing the principle that the fairness of the death penalty demands particularized consideration of specific individuals and their crimes. The former enables a court to engage in such particularized consideration, but, in doing so, renders it difficult, if not impossible, to generate the comparisons that assure us that any given death sentence is not arbitrary.

C. Identifying a Test of Proportionality

Once a court has identified the universe of cases to be considered and specified the pool of cases deemed similar to the current case, it must then resolve the final step in the logic of comparative proportionality review: determination of whether a case under review is (or is not) proportionate to others in the pool it has defined. Methods of answering this question range from the intuitive to the statistical, and all suffer from serious flaws.141

141 See USER MANUAL, supra note 123, at A-10 to A-11 (providing recommendations of the National Center for State Courts on this question). See generally BALDUS, EQUAL JUSTICE, supra note 113 (analyzing the effectiveness of different methods of conducting comparative

sentenced to die between 1973 and 1999, only 598 were executed).
Conduct of Comparative Proportionality Review

The least systematic approaches to this question involve an appeal to some unarticulated notion of “reasonableness” in defending the conclusion that a specific death sentence is or is not proportionate, 142 or, alternatively, a precedent-seeking approach. 143 The former is unacceptable because, in employing it, courts fail to specify the criteria that inform their judgments, thereby rendering their decisions undefended and uncontestable. 144 Using the latter approach, regardless of whether similarity is defined in terms of factual congruence or like aggravating factors, courts attempt to discover one or more similar cases in which the death penalty was previously imposed and upheld (ignoring the difficulties noted above in specifying the criteria of similarity). 145 Arguably, this approach encourages ad hoc assessments based on unarticulated or insufficiently articulated notions of similarity, and, as such, does little if anything to meet Furman’s concerns about patterns of inconsistent death sentencing. 146 Moreover, because this sort of inquiry presupposes restriction of the pool of comparable cases to those in which sentences of death were imposed and upheld, it is biased in favor of affirmation. This predisposition proves still more troublesome when one considers that the affirmed cases cited in order to justify a death sentence may themselves have been arbitrarily imposed. When that is so, a precedent-seeking method will simply perpetuate and exacerbate the injurious effects of a chain of problematic sentences, as the most proportionality review); Bienen, supra note 102 (criticizing the conduct of comparative proportionality review on the ground that it provides only an illusion of fairness); Hartman, supra note 113 (assessing Pennsylvania’s analysis of comparative proportionality review); Liebman, supra note 113 (assessing Georgia’s analysis of comparative proportionality review); Lustberg & Lapidus, supra note 113 (assessing New Jersey’s analysis of comparative proportionality review); Rodriguez, supra note 113 (arguing that New Jersey’s conduct of comparative); Wallace & Sorenson, Explanations, supra note 65 (assessing Missouri’s analysis of comparative proportionality review).


143. See generally BALDUS, EQUAL JUSTICE, supra note 113 (discussing precedent-seeking approaches to the conduct of comparative proportionality review).

144. USER MANUAL, supra note 123, at A-5 to A-6 (criticizing an “intuitive approach” on the ground that it renders courts unable to demonstrate that their reasoning is consistent and fair across cases).

145. See Baldus, Georgia Experience, supra note 133 at 669–70 (criticizing the precedent-seeking approach because courts are “satisfied upon finding one or two prior cases, the circumstances of which make them suitable benchmarks for the death sentence on appeal,” which is insufficient to determine whether the death penalty is generally applied in comparable cases).

146. See Baldus, Georgia Experience, supra note 133, at 718–20.
recently affirmed now becomes available for citation as precedent in future cases.

Because these first two approaches rely on either intuitive notions of reasonableness or on identification of a handful of comparable precedents, they appear inconsistent with Gregg. In conducting comparative proportionality review, the Gregg plurality stated, courts should ask whether juries do or do not “generally . . . impose the death sentence in a certain kind of murder case.”147 This exhortation suggests a third method of proportionality analysis, involving an assessment of the frequency with which capital sentences are imposed among cases deemed similar. Unlike the precedent-seeking approach, this requires that courts consider not merely death-sentenced cases, but also those similar cases that resulted in some other sentence, most typically life imprisonment. Only then can a court determine whether imposition of the death penalty in a case on review is comparatively excessive because this penalty is imposed so infrequently in the larger class of cases to which it is similar.

The basic premise of frequency analysis is that the greater the frequency of death sentences within any given comparison group, the more confident a court can be in adjudging any given sentence in that group to be proportionate.148 The effort to calculate such frequencies is not without its problems, however. For example, in order to evaluate a defendant’s character, a review must take into account mitigating factors. However, for reasons indicated above, any attempt to categorize such factors for the purpose of determining their relative frequency appears to be a fool’s errand. Specification of how to take mitigating factors into account is additionally complicated by the fact that juries need not be unanimous when determining mitigation and because they need not elaborate their reasons for imposing or refusing to impose a death sentence.149 As such, the commitment to frequency analysis of the mitigating factors that account for a jury’s determination represents an attempt to quantify that which is most often unknown and that even if known does not lend itself to quantification. More generally, as the example of mitigating factors suggests, the conduct of frequency analysis draws a court away from careful consideration of the qualitatively distinct features of prior cases, and so compromises its

149. See Latzer, supra note 51, at 1220.
Conduct of Comparative Proportionality Review

effort to give full weight to the unique characteristics of defendants, their crimes, and their respective degrees of culpability.

Leaving aside the complex methodological questions regarding just how frequency analysis should be conducted, a court committed to this method must decide what rate of death sentence imposition is sufficient to sustain the conclusion that a sentence under review is not disproportionate. Just what that percentage should be is susceptible to disagreement, depending in large measure on what a court takes to be the principal purpose of comparative proportionality review. If the primary purpose of such review is to ensure that patently aberrant sentences are set aside, then a low threshold of frequency (e.g., 33%) in similar cases could be considered sufficient to guarantee capital punishment’s fairness. Alternatively, if the primary purpose of such review is to ensure that death sentences are only upheld when, in similar cases, this punishment is “generally” imposed, then arguably a much higher threshold (e.g., 75%) is required.

In a similar vein, a commitment to frequency analysis does not in and of itself specify what percentage of similar cases is required in order to accomplish retribution and deterrence, capital punishment’s two constitutionally permissible rationales. Determination of that rate, moreover, may well vary if one considers one of these purposes more central than the other. For example, if one believes that capital punishment is justified chiefly as a deterrent, as Justice White’s opinion in Furman suggests, then the death penalty must be imposed sufficiently often in specified categories of murder to deter effectively. That in turn may suggest that findings of disproportionality should be confined to cases that are grossly out of line with those previously appealed. However, if one believes that capital punishment is justified chiefly in retributive terms, as Justice Stewart’s opinion in Furman suggests, then the frequency of capital punishment’s imposition becomes a less significant concern, and so a court may conclude that a lower threshold is sufficient to justify a finding of disproportionality.

150. Because my aim in this section is to outline the basic logic of comparative proportionality review, these methodological questions, while crucial for some purposes, are not essential to my argument. See generally David C. Baldus et al., Identifying Comparatively Excessive Sentences of Death: A Quantitative Approach, 33 Stan. L. Rev. 1 (1980) (reviewing the statistical methods appropriate to this task).

151. See Gregg, 428 U.S. at 183 (plurality opinion).


153. Id. at 308 (Stewart, J., concurring).
Finally, a commitment to frequency analysis may render it necessary for a court to choose between competing claims of fairness. On the one hand, in order to avoid the charge of arbitrariness, a court may specify a precise criterion of frequency that warrants imposition of the death penalty on the members of a group of defendants whose character and crimes are deemed comparable. In doing so, however, a court opens itself to the criticism that it is employing an overly formalistic procedure that renders it impossible to take into adequate account the idiosyncratic features that may render this particular defendant less worthy of death than another, and so denies the constitutional imperative of individualized sentencing in death penalty cases. 154 On the other hand, a court may refuse to specify a settled frequency standard so as to ensure its ability to engage in individualized sentencing. In doing so, however, it is susceptible to the charge of arbitrariness on the ground that it is then free to select among different frequencies in order to ratify results determined on other and possibly undeclared bases.

III. COMPARATIVE PROPORTIONALITY REVIEW IN WASHINGTON

The current Washington statute dealing with aggravated first-degree murder and capital punishment requires the Washington State Supreme Court to review all death sentences imposed in the state. 155 On review, the court must determine whether there was “sufficient evidence” to justify the jury’s determination that leniency was not warranted; 156 whether the death sentence was “brought about through passion or prejudice” 157 whether the defendant was “mentally retarded” 158 and, finally, of principal interest here, “[w]hether the sentence of death is excessive or disproportionate to the penalty imposed in similar cases, considering both the crime and the defendant.” 159 Should the court answer the first of these questions in the negative, or any of the remaining three questions in the affirmative, the death sentence under review shall be invalidated and the case remanded to the trial court for

156. Id. § 10.95.130(2)(a).
157. Id. § 10.95.130(2)(c).
158. Id. § 10.95.130(2)(d).
159. Id. § 10.95.130(2)(b).
Conduct of Comparative Proportionality Review

This statute, which became effective in 1981, is the fruit of a complex political struggle that was provoked by the U.S. Supreme Court’s ruling in *Furman*. In 1972, the U.S. Supreme Court vacated a Washington death sentence based on *Furman*. Three years later, the Washington State Legislature abolished the death penalty altogether. However, in the November general election of 1975, Washington voters approved Initiative Measure No. 316, which required imposition of a death sentence in all cases of aggravated first-degree murder. Washington State Attorney General Slade Gorton declared that initiative unenforceable based on *Woodson*, which invalidated mandatory death sentence statutes on the ground that they prevented juries from taking into account the individualized character and circumstances of particular defendants.

Following *Gregg*, Gorton recommended that any new capital punishment statute adopted in Washington follow the model of Georgia. Specifically, he suggested that the statute provide for a bifurcated trial, include a mandatory appeal to the State Supreme Court, and require that court to conduct comparative proportionality reviews in order to ensure “that the death penalty imposed is consistent with other sentences imposed in other trials under similar circumstances.” Following receipt of these recommendations, during the 1977 legislative session Representative Earl Tilly drafted and introduced House Bill 615, which was closely patterned after the Georgia statute and incorporated verbatim its comparative proportionality provision. The final version of that bill, which retained this provision without alteration, became law on June 3, 1977.

The 1977 statute required the clerk of the trial court, within ten days...

---

160. Id. § 10.95.140(1).
168. Id. at 13.
of receipt of the transcript of any trial culminating in a death sentence, to transmit that record to the State Supreme Court, along with a notice providing basic information about the case as well as “a report prepared by the trial judge . . . in the form of a standard questionnaire prepared and supplied by the supreme court of Washington.”\textsuperscript{171} In addition, the statute authorized the court to affirm the death sentence or remand the case for resentencing by the trial court judge, who is to be furnished with “[t]he records of those similar cases referred to by the supreme court of Washington in its decision and the extracts prepared therefor.”\textsuperscript{172} The 1977 statute, however, did not furnish any additional guidance on either the information to be collected via the required questionnaire or the use of the reports generated by that information by the Washington State Supreme Court.\textsuperscript{173} The statute also did not indicate the universe of cases to be considered for the purpose of conducting comparative proportionality review.\textsuperscript{174} Moreover, the statute did not specify the criteria to be employed in determining which cases are similar, or how to determine whether any given sentence is or is not “excessive” or “disproportionate.”\textsuperscript{175} Finally, the statute did not specify exactly what a trial judge was to do on remand with the records of similar cases or, for that matter, what sentence to impose if the State Supreme Court found that a given death sentence was disproportionate.\textsuperscript{176}

One year later, in 1978, in response to a request by Chief Justice Wright, and taking the Georgia equivalent as her prototype, a commissioner of the Washington State Supreme Court formulated a draft questionnaire to be completed by trial court judges.\textsuperscript{177} After approving temporary use of this initial version,\textsuperscript{178} the court established a Task Force on the Death Penalty Questionnaire to revise that draft.\textsuperscript{179}

\begin{itemize}
  \item \textsuperscript{171} Id. § 7(1).
  \item \textsuperscript{172} Id. § 7(5)(b).
  \item \textsuperscript{173} See id. § 7.
  \item \textsuperscript{174} See id. However, by requiring the submission of trial records and questionnaires only in death penalty cases, the statute arguably implied that the court’s comparison would be similarly restricted. See id. § 7(1).
  \item \textsuperscript{175} Id. § 7(3)(b).
  \item \textsuperscript{176} Id. § 7.
  \item \textsuperscript{177} Memorandum from Joan Smith Lawrence, Commissioner, to the Washington State Supreme Court (Feb. 15, 1978).
  \item \textsuperscript{179} The task force was created in April 1978 by the Washington State Supreme Court via its Judicial Council. Its purpose, according to minutes of its initial meeting, was “to assist the Supreme
Conduct of Comparative Proportionality Review

This group met from May 1978 to February 1979, at which time it proposed a final version of this questionnaire as well as a rule to implement its use.180

The minutes of the task force indicate a marked gap between its expectations with regard to the purpose of the trial judge reports and the subsequent practice of the Washington State Supreme Court in conducting comparative proportionality review.181 The judges and attorneys who participated in that task force repeatedly affirmed their belief that the results generated by these questionnaires were to play a central role in informing the court’s performance of its statutorily mandated reviews: “[t]he purpose of the [trial] report is to aid the Supreme Court in its review of a death sentence by providing the Court with all possible information regarding the defendant and the proceedings, particularly by eliciting from the trial judge his own unique perspective of the trial.”182 Furthermore, the members agreed that “such a perspective could be most effectively communicated by a report that was less a checklist of information also appearing in the record and more a report that included general questions calling for analysis, by the trial judge, of any extraneous factors not appearing in the record that may have influenced the jury’s verdict.”183 Finally, the members expressed their collective view that “[b]ecause the conduct of this report has the potential of influencing a decision upon which a defendant’s life depends, it is imperative that it be properly completed, without oversight. A conscientious trial judge will ensure that the report is carefully completed and factually correct,”184 and that it be completed in

180. Mins., Task Force on the Death Penalty Questionnaire (Feb. 1, 1979). Not long after the task force submitted its recommendation, the 1977 capital punishment statute was found unconstitutional as a result of Washington State Supreme Court rulings in 1980 and 1981. See State v. Frampton, 95 Wash. 2d 469, 478–79, 627 P.2d 922, 926–27 (1981); State v. Martin, 94 Wash. 2d 1, 2, 614 P.2d 164, 164 (1980). As a result, the task force’s questionnaire was never formally adopted by the State Supreme Court or employed by trial court judges. The adoption of this questionnaire would not have required legislative approval because the 1977 statute mandated that the State Supreme Court prepare and supply such a form. See Death Penalty—Aggravated Murder, ch. 206, § 7(3)(b), 1977 Wash. Laws 778. With relatively minor alteration, the text of the task force’s proposed questionnaire was incorporated into the aggravated first-degree murder and capital punishment statute in 1981. Murder, Sentencing, ch. 138, 1981 Wash. Laws 535, 535–47.

181. See infra Part IV.B.


183. Id.

184. Rough Draft Questionnaire (attached to Agenda, Task Force on the Death Penalty
a timely fashion, defined by the task force as thirty days after sentencing.185

Of equal significance, in considering the appropriate universe of cases to be examined for the purposes of comparative proportionality review, and departing from the clear implication of the 1977 statute, the task force recommended that this questionnaire be

required in every case in which a defendant is charged with, and convicted of, murder in the first degree . . . and a special sentencing proceeding is held . . . . The report is required regardless of whether the death penalty is imposed at the conclusion of the sentencing proceeding and regardless of whether the defendant pursues an appeal.186

Justifying this recommendation, which was eventually incorporated into its proposed criminal rule for superior courts, the task force concluded that:

this approach was necessary to furnish the Supreme Court with the data it needs to compare the propriety of the death sentence in a particular case with the propriety of sentences in other cases. The task force did not feel it would be appropriate to compare a death penalty case only to other cases in which the death penalty was actually imposed.187

Following discussion and revision of multiple drafts, the task force recommended to the Washington State Supreme Court a questionnaire that is similar and, in many respects, identical to the one currently employed.188 Both are divided into multiple sections, and both request information about the chronology of the case; the defendant; the trial; the special sentencing proceeding, if conducted; the victim; the legal representation provided to the defendant; and, finally, “general considerations,” which deals with the race, ethnic, and sexual orientation of the various participants in the trial, including the jury, as well as the

185. Proposed Criminal Rule for Superior Court 1 (attached to Letter from Karl B. Tegland, Staff Attorney, Washington Judicial Council, to John J. Champagne, Clerk, Washington Supreme Court (Feb. 1, 1979)) (citation omitted). The task force allotted an additional ten days to the prosecution and defense counsel to offer comments prior to final submission of the report.

186. Id.


Conduct of Comparative Proportionality Review

demographics of the county in which the trial was conducted. 189 Included in the task force recommendation, but excluded from the form currently in use, is a question asking the trial court judge to “state anything not evoked by this questionnaire which may have affected the trial or the sentence imposed, or which you believe should be brought to the attention of the Supreme Court.” 190 That question has been replaced in the current form by a comprehensive request for “[g]eneral comments of the trial judge concerning the appropriateness of the sentence, considering the crime, the defendant, and other relevant factors.” 191

Not long after the task force submitted its recommendation, the 1977 capital punishment statute was found unconstitutional as a result of 1980 and 1981 Washington State Supreme Court rulings, each of which held that this law chilled a defendant’s right to trial by permitting those who pled guilty to avoid a sentence of death. 192 A new capital punishment proposal, House Bill 76, was introduced in the legislature in 1981. 193 Following a series of amendments, 194 it became law on May 14, 1981. 195 Four features of that bill’s initial formulation and subsequent revision are crucial to an understanding of the present conduct of comparative proportionality review in Washington.

189. See Mins. (Feb. 1, 1979), supra note 180.
191. See App. A, infra p. 880, at (6)(k). Also included in the task force recommendation, but excluded from the form currently in use, is a question asking the judge to comment on the conduct of the defendant in the presence and absence of the jury as well as “any other significant conduct, characteristics, or other factors concerning the defendant.” See Proposed Criminal Rule for Superior Court 10 (attached to Letter from Karl B. Tegland, Staff Attorney, Washington Judicial Council, to John J. Champagne, Clerk, Washington State Supreme Court (Feb. 1, 1979)). In addition, excluded from the form currently in use is a question aimed at ensuring that the report be submitted to prosecution and defense counsel, each of whom were then to be given an opportunity to propose corrections of factual data. See id. at 1. This opportunity is effectively afforded to counsel by the statute currently in force, which invites both sides to “submit briefs within the time prescribed by the court and present oral argument to the court.” WASH. REV. CODE § 10.95.130(1) (2000).
192. In 1980, the Washington State Supreme Court ruled that the 1977 statute did not allow for imposition of a death sentence if a defendant pled guilty to aggravated first-degree murder. State v. Martin, 94 Wash. 2d 1, 2, 614 P.2d 164, 164 (1980). One year later, in light of Martin, the court found the 1977 capital punishment statute unconstitutional because it permitted a defendant to escape the death penalty by pleading guilty, whereas those who elected to contest their guilt remained candidates for death, thus impermissibly chilling the former’s right to trial. State v. Frampton, 95 Wash. 2d 469, 478–79, 627 P.2d 922, 926–27 (1981).
First, unlike the 1977 statute, which did not specify the content of the trial judge report, the 1981 bill enumerated each of the specific questions to be answered as well as the factual information to be provided by the trial judge.196 This section was incorporated into the law enacted later that same year without amendment.197 As the explanatory memorandum introduced in conjunction with this bill stated, the section of the 1981 bill that specified the content of the trial judge report was in large part derived from the form submitted by the task force two years prior.198 By codifying that questionnaire, the bill’s sponsors arguably endorsed the task force’s understanding of its purpose and role in State Supreme Court reviews of death sentences.

Second, the original version of the 1981 bill, as in the 1977 statute, required that trial judge reports be filed only in cases in which a defendant was sentenced to death.199 This provision, however, was eventually rejected in favor of a requirement that such reports be filed “[i]n all cases in which a person is convicted of aggravated first-degree murder.”200 Thus, by expanding the universe of cases, the legislature indicated its appreciation of one of the key prerequisites of meaningful comparative proportionality review. Specifically, this provision expressed the legislature’s belief that the State Supreme Court’s ability to determine why the death penalty is imposed in a few cases but not in most is impossible if it considers only those aggravated first-degree murder cases that culminate in capital sentences.201

Third, as initially formulated, the 1981 bill defined “similar cases” as those “reported in the Washington Reports or Washington Appellate Reports since January 1, 1965.”202 This language notably omits reference to the trial judge reports, and had it been adopted, it is likely that review of specific cases would not have included examination of the reports. This reading is borne out by the explanatory memo: “Frankly, the [trial judge] report is of marginal value since the Supreme Court in its

198. Explanatory Material for “An Act Concerning Murder and Capital Punishment,” Dec. 31, 1980, at 18 (“The form of the report contained in this proposal [HB 76] is largely the product of a task force which was appointed by the supreme court to develop a form under our current statute. There are, of course, modifications to accommodate this revised capital murder scheme.”).
201. See id.
Conduct of Comparative Proportionality Review

sentencing review will examine the entire record of the trial."\(^{203}\) However, the memo did acknowledge that comparative proportionality review is “an important feature in a sentencing review for it will enhance uniformity in the imposition of the death penalty.”\(^{204}\) The bill’s text was changed in the final law, which reads as follows:

For the purposes of this subsection, “similar cases” means cases reported in the Washington Reports or Washington Appellate Reports since January 1, 1965, in which the judge or jury considered the imposition of capital punishment regardless of whether it was imposed or executed, and cases in which reports have been filed with the supreme court under RCW 10.95.120.\(^{205}\)

With this last clause, which references the section of the statute that mandates submission of trial judge reports following all convictions for aggravated first-degree murder, the legislature expressly required that the cases covered by these reports be incorporated into the process of comparative proportionality review.\(^{206}\)

Fourth, and finally, the 1981 bill clarified an ambiguity of the 1977 statute by specifying what was to happen should a death sentence be vacated and then remanded “for resentencing by the trial judge based on the record and argument of counsel.”\(^{207}\) At least in principle, the 1977 statute left open the possibility that, on remand, a death sentence might be reimposed by the trial court judge. The 1981 bill, however, stated that should the State Supreme Court conclude that a death sentence was wrongly imposed for any of the reasons enumerated in the statute, including a finding of disproportionality, the court “shall invalidate the sentence of death.”\(^{208}\) More precisely, as the explanatory memorandum indicated, should any given death sentence be found invalid, “at the re-sentencing the defendant would get life without parole.”\(^{209}\)

One year later, the Washington State Supreme Court affirmed the constitutionality of this statute in *State v. Bartholomew.*\(^{210}\) The court

---

206. See *id*.
reversed Bartholomew’s sentence on the ground that the 1981 statute failed to limit the evidence that the prosecution may present at the sentencing phase of capital proceedings.211 It nevertheless refused to strike down the statute in its entirety, holding that the constitutional infirmity could be remedied by application of the severability provision of the Revised Code of Washington (RCW) section 10.95.900.212 More important for this analysis, when the court considered the adequacy of the procedural protections devised by the state legislature in order to meet the concerns articulated in Furman, it endorsed them, and in doing so expressly cited, among other protections, the “elaborate automatic review procedure which brings it under the scrutiny of this court.”213 As the court noted in subsequent cases involving such review, that process must incorporate a review of the trial judge reports in order to determine whether any given sentence falls afoul of the requirements of proportionality (although, as indicated in the following Part, its practice falls far short of this commitment).214

IV. THE DEFICIENCIES OF COMPARATIVE PROPORTIONALITY REVIEW IN WASHINGTON

This Part offers two criticisms of the practice of comparative proportionality review in the State of Washington. Section A explores deficiencies in the trial judge reports, which furnish the data regarding the universe of cases that the court is to consider in conducting its reviews. Section B examines the way in which the court has actually proceeded in undertaking such reviews. Section B suggests that the court is inadequately performing its statutorily mandated reviews. Section A suggests that, even if it were to attempt to remedy the defects of its current practice, it would be unable to do so, given the deficiencies of the data available to it.

211. Id. at 195–96, 654 P.2d at 1183–84.
212. Id. at 176, 654 P.2d at 1173.
213. Id. at 192, 654 P.2d at 1182.
214. See State v. Gentry, 125 Wash. 2d 570, 655–56, 888 P.2d 1105, 1154–55 (1995) (reviewing the trial judge reports to determine presence of a pattern of imposition of the death penalty based on race of defendant or victim); see also State v. Benn, 120 Wash. 2d 631, 680–93, 845 P.2d 289, 317–24 (1991) (reviewing the trial judge reports to determine whether a sentence was disproportionate); State v. Lord, 117 Wash. 2d 829, 907, 910–11, 822 P.2d 177, 221, 223 (1991) (concluding that, in order to determine that a sentence is neither arbitrary nor based on illegitimate racial factors, a court should review the universe of cases set out in RCW 10.95.120).
Conduct of Comparative Proportionality Review

A. Washington’s Trial Judge Reports and Their Defects

Comparative proportionality review was originally developed in order to remedy two problems identified in Furman: first, the inconsistent and arbitrary application of the death penalty, and hence the absence of any rational basis for distinguishing between those who are sentenced to death and those who are not; and, second, the concern that constitutionally impermissible factors, such as race and economic status, may in fact offer the most salient explanation for this distinction.\(^\text{215}\) The capacity of the Washington State Supreme Court to perform such review in a meaningful way depends in large measure on the completeness and accuracy of the information that is included in the trial judge reports. As David Baldus, the preeminent student of comparative proportionality review, notes: “The most important role for state courts is to develop a database that provides an overview of the system and reliable information on the universe of cases that are used as comparison cases in individual reviews.”\(^\text{216}\) Yet an analysis of the trial judge reports submitted to the Washington State Supreme Court indicates a host of problems which critically compromise the court’s capacity to fulfill its statutory mandate.

This analysis includes all 259 trial judge reports filed with the Washington State Supreme Court, beginning with those first submitted in response to adoption of the 1981 capital punishment statute and running through March 2003.\(^\text{217}\) Rather than offering an exhaustive


\(^\text{216}\) David C. Baldus, When Symbols Clash: Reflections on the Future of Comparative Proportionality Review of Death Sentences, 26 SETON HALL L. REV. 1582, 1598 (1996); cf. Bienen, supra note 102, at 164 (explaining that the statutory mandate to conduct proportionality review necessarily implies creating and maintaining an adequate factual record).

\(^\text{217}\) The most recent trial judge report considered in this study, filed with the Washington State Supreme Court as of March 24, 2003, is numbered 255. The discrepancy between that number and the number of trial judge reports submitted since 1981 is explained as follows: (1) Reports of the Trial Judge Nos. 16/16A concern two different trials for different defendants and incidents; (2) Reports of the Trial Judge Nos. 34/34A concern two separate trials for different incidents with the same defendant; (3) Reports of the Trial Judge Nos. 85/85A and 97/97A each represent two separate reports for the same defendant and incident. Although RCW 10.95.120 requires the State Supreme Court to consider “cases reported in the Washington Reports or Washington Appellate Reports since January 1, 1965,” in addition to those for which trial reports have been filed, I have not considered the pre-1981 cases because that would entail consideration of convictions and sentences rendered under an unconstitutional statute. See Bruce Gilbert, Comparative Proportionality Review: Will the Ends, Will the Means, 18 SEATTLE U. L. REV. 593, 621 (1995). Because the pre-1981 statute was found unconstitutional, Gilbert argues, it may have generated death sentences that were influenced by discrimination or other impermissible considerations, which, in turn, could provide a
account of the deficiencies of the trial judge reports, this section highlights those that are most troubling in light of the principal purposes of comparative proportionality review. Although the deficiencies are interrelated, for the sake of clarity, this section divides these problems into four categories: (1) reports that are absent, unrevised, and/or late; (2) reports that fail to provide accurate and/or adequate information regarding defendants and victim(s); (3) reports that fail to provide accurate and/or adequate information regarding aggravating and mitigating factors; and (4) reports that fail to provide accurate and/or adequate information regarding the racial and ethnic identities of various participants in capital trials.

1. Absent, Unrevised, and Late Reports

Perhaps the most obvious problem with the Washington State Supreme Court’s database is its omission of cases that should have been included. A comparison of the trial judge reports on file with the Washington State Supreme Court with the records compiled by the Washington Sentencing Guidelines Commission, which maintains data on criminal judgments and sentences entered by Washington trial courts, reveals that the court is missing reports for twelve aggravated constitutionally flawed foundation for finding a contemporary case proportional. See id. In Benn, Justice Utter stated:

Admittedly, the cases between 1965 and 1981 are less useful in many respects for conducting proportionality review. In making comparisons, we do not have the benefit of the more detailed information contained in the questionnaires trial courts have completed since 1981 as required by RCW 10.95.120. In addition, defendants in those cases were sentenced under earlier versions of Washington’s death penalty statute which have subsequently been declared unconstitutional. Benn, 120 Wash. 2d at 705, 845 P.2d at 330 (Utter, J., dissenting). While this is a persuasive justification for not considering pre-1981 cases, there still remains a significant contradiction between what RCW 10.95.120 formally requires in terms of its definition of “similar” cases and the universe of cases actually considered by the State Supreme Court in conducting comparative proportionality reviews.


The Washington State Legislature enacted the Sentencing Reform Act (“SRA”), which established the Sentencing Guidelines Commission and directed it to recommend to the Legislature a determinate sentencing system for adult felonies. The principal goal of the new sentencing guidelines system was to ensure that offenders who commit similar crimes and have similar criminal histories receive equivalent sentences. Sentences were to be determined by the seriousness of the offense and by the criminal record of the offender. The Commission completed the original adult felony sentencing ‘grid’ in 1982, and the Legislature enacted it
Conduct of Comparative Proportionality Review

first-degree homicide convictions.219

An inspection of the records kept by the Washington State Department of Corrections reveals an additional missing record.220

Leaving aside the fact that these thirteen missing records evidences a violation of the statutory requirement that trial judge reports be submitted not in some, but “in all cases in which a person is convicted of aggravated first-degree murder,”221 absent these reports, any attempt to

into law in 1983. The Sentencing Reform Act took effect for crimes committed on and after July 1, 1984. Codified in chapter 9.94A RCW, the SRA contains the guidelines and procedures used by the courts to impose sentences for adult felonies.


219. The first case included in the database compiled by the Sentencing Guidelines Commission was filed in 1985 (Robert Strandy, Cause No. 85-1-00487-6 (1985)). It is not unreasonable to assume that were the Commission’s records to extend back to 1981, when the trial judge reports currently in use were first required by law, the number of reports missing from the State Supreme Court collection would be larger. In chronological order by date of sentence, convictions for aggravated first-degree murder on record with the Sentencing Guidelines Commission, but missing from the trial judge reports, are as follows: Michel McBride, Cause No. 86-1-00523-5 (1986); Daniel Edwards, Cause No. 87-1-00550-1 (1987); Billy Ballard, Cause No. 91-1-00083-5 (1991); Miguel Gaitan, Cause No. 93-1-01018-0 (1993); Robert Anderson, Cause No. 97-1-00128-0 (1997); Chad Walton, Cause No. 97-1-02153-1 (1998); Brandon Backstrom, Cause No. 97-1-01993-6 (1999); Alex Baranyi, Cause No. 97-1-00343-8 (1999); Michael Thorton, Cause No. 98-1-00493-6 (1999); Kevin Cruz, Cause No. 00-1-00284-6 (2002); Christopher Miller, Cause No. 01-1-00311-3 (2002); Donald Durga, Cause No. 01-1-011709-1 (2002). In addition, at least ten defendants present in the State Supreme Court’s database are missing from that maintained by the Sentencing Guidelines Commission. By trial judge report number, they are as follows: Susan Kroll (No. 60); Martin Lee Sanders (No. 81); Stanley Runion (No. 99); Constantine Baruso (No. 112); Tommy Metcalf (No. 128); Cal Brown (No. 140); Darold Stenson (No. 144); William Robinson, Jr. (No. 166); Dwayne Antony Woods (No. 177); Dayva Cross (No. 220).

220. No trial judge report has been filed for David Anderson, Cause No. 97-1-00421-3 (2000). There is some confusion regarding this case because, according to the Department of Corrections, the name that corresponds to this cause number is Richard Hampton. A CourtLink search indicates that the name corresponding to this number is in fact David Anderson, who was convicted by a jury on four counts of aggravated first-degree murder in 1997 and sentenced to life imprisonment without parole.

221. WASH. REV. CODE § 10.95.120 (2000). In a brief filed with the Washington State Supreme Court, the prosecuting attorney contended that the “absence of a couple dozen reports” does not pose a legal problem because RCW 10.95.130(2)(b) only requires the court to consider “cases in which reports have been filed with the supreme court under RCW 10.95.120.” Brief for Respondent at 176–77, Thomas v. State, 150 Wash. 2d 821, 83 P.3d 970 (2004) (No. 70727-8). On this reading, the court is required to consider only those reports that have actually been filed, not those that are mandated by law. Surprisingly, given this view, the brief proceeds to acknowledge that “the court should try to ensure that all of the mandated reports get filed.” Id. (emphasis omitted). Then, citing the example of New Jersey, it recommends appointment of a retired judge to the position of standing master in order “to supervise . . . the court’s data collection as well as the completion of the trial judge questionnaires.” Id. at 178. Finally, it states that “[t]he collection of accurate data from all aggravated murders (whether or not the death penalty is imposed) is imperative for
identify the full range of cases similar to that on review, whether by seeking those with analogous fact patterns or comparable aggravating factors, will be incomplete. Moreover, the absence of these reports from the supreme court’s database will frustrate any attempt to determine the frequency with which convictions for aggravated first-degree murder, in cases deemed similar to that on review, generate a sentence of death as opposed to life without possibility of parole. Still more troublesome, all of the defendants in the thirteen cases missing from the supreme court’s database were sentenced to life without parole. Should the court seek to determine that frequency, the absence of these reports will skew its calculations in favor of death.

In addition, the database maintained by the State Supreme Court includes cases that either ought not to be in the database or, at the very least, ought not to be present absent some indication of their current status. This includes fifteen trial judge reports (17.6% of the cases in which the death penalty was sought) for defendants who, although initially sentenced to death, have since been re-sentenced to life without parole, as well as those who have had their convictions and death sentences set aside and are no longer on death row. The failure to proportionality review.” Id. at 179. Leaving aside what appears to be a direct self-contradiction, this argument fails to explain why such an extraordinary measure as the appointment of a special master is necessary in light of its contention that the law’s mandate is satisfied if the court merely considers those reports that are indeed filed. In addition, the brief contends that the defendant has failed to explain why the absence of some unspecified number of trial judge reports “prevents” proportionality review. See id. at 177. Granted, the absence of any number of trial judge reports does not render the conduct of comparative proportionality review literally impossible. The larger the number of missing reports, however, the less adequate and meaningful such review will be. That is so regardless of whether the court engages in a quantitative effort to calculate the frequency with which “similar” aggravated first-degree murders do or do not result in death sentences, or if it seeks to engage in a qualitative review that compares the facts of any given case on review to cases deemed similar.

222. See supra note 219.


The proportionality database is itself flawed because it . . . is missing a staggering number of cases required by RCW 10.95.130 in which defendants were convicted of aggravated first-degree murder, but no death penalty was imposed. Even if the court were to consider whether death was generally imposed in similar cases (which it no longer apparently does), the universe of cases we are supposed to consider is markedly skewed by these factors against the defendant, in favor of imposing death, and contrary to the statute.

Id. (emphasis in original). While the number of missing trial judge reports is perhaps not “staggering,” as Justice Sanders suggested, his basic point remains valid.

224. Cases involving defendants once sentenced to death but now re-sentenced to life without parole include the following: Lord v. Wood, 184 F.3d. 1083 (9th Cir. 1999); Jeffries v. Wood, 114 F.3d 1484 (9th Cir. 1997); Rupe v. Wood, 863 F. Supp. 1315 (W.D. Wash. 1994), aff’d, 93 F.3d
Conduct of Comparative Proportionality Review

revise these reports in order to reflect changes in the status of these defendants introduces a significant inaccuracy because, as a matter of law, these death sentences no longer stand. One might maintain, as the


A related complication is posed by cases like that of Timothy Cronin. See State v. Cronin, 142 Wash. 2d 568, 14 P.3d 752 (2000). In 1997, Cronin was convicted of aggravated first-degree murder and sentenced to life imprisonment without parole. Id. at 577, 14 P.3d at 757. The Washington State Supreme Court subsequently overturned his conviction due to errors in the instructions given to the jury. Id. at 581–82, 14 P.3d at 758–59. However, the court also concluded that because Cronin's first-degree felony murder conviction was unaffected by the instructional error, his felony murder conviction should be affirmed. Id. at 586, 14 P.3d at 761. Because Cronin was convicted of aggravated first-degree murder, arguably the law requires that a trial judge report be filed with the supreme court, as it was. Report of the Trial Judge No. 217 (Timothy Cronin). However, once his conviction on that charge was overturned, this report should either be removed from the court's database, or it should be amended to reflect its present status. Neither step, however, has been taken.

225. The question of whether juvenile defendants should be included in the State Supreme Court’s database is more complicated. In Stanford v. Kentucky, 492 U.S. 361 (1989), the U.S. Supreme Court upheld imposition of the death penalty for defendants who were sixteen or seventeen when they committed their crimes on the grounds that such imposition does not constitute a violation of the Eighth Amendment’s Cruel and Unusual Punishment Clause. Id. at 380. However, one year earlier, in Thompson v. Oklahoma, 487 U.S. 815 (1988), a plurality had ruled that the death penalty cannot be imposed on defendants who were fifteen or younger when they committed the crime on the grounds that a death sentence serves no valid retributive or deterrent purpose in such cases. Id. at 836–37 (plurality opinion). In her concurring opinion, Justice O’Connor concluded that defendants who were under sixteen when they committed their crimes “may not be executed under the authority of a capital punishment statute that specifies no minimum age at which the commission of a capital crime can lead to the offender’s execution.” Id. at 857–58 (O’Connor, J., concurring). In light of this claim, the Washington State Supreme Court, in State v. Furman, 122 Wash. 2d 440, 858 P.2d 1092 (1993), concluded that, because the state capital punishment statute did not expressly provide for imposition of the death penalty on minors and did not specify any minimum age for its imposition, those under eighteen could not be sentenced to death. Id. at 457–58, 858 P.2d at 1102–03. Included in the State Supreme Court’s database, however, are twenty-five defendants who were juveniles at the time they committed the crimes for which they were eventually convicted on a charge of aggravated first-degree murder. In three of these cases, the death penalty was sought, and in Michael Furman’s case, it was imposed (but later reversed). Id. at 443, 858 P.2d at 1095. On the one hand, because Washington State’s capital punishment statute
Washington State Supreme Court has done,\textsuperscript{226} that these cases should remain in the database on the ground that the statute requires inclusion of all cases in which “the judge or jury considered the imposition of capital punishment regardless of whether it was imposed or executed.”\textsuperscript{227} However, if these reports are not updated,\textsuperscript{228} and if the supreme court seeks to calculate the frequency with which convictions for aggravated first-degree murder generate a sentence of death as opposed to life without parole, this error will once again skew the results in favor of

\begin{quote}

clearly requires trial court judges to complete reports for all defendants convicted of aggravated first-degree murder, reports regarding juvenile defendants must be submitted and maintained in the State Supreme Court’s database. On the other hand, because RCW 10.95.130(2)(b) defines “similar cases” as those in which “the judge or jury considered the imposition of capital punishment,” and because this punishment cannot be considered in the case of juvenile defendants, arguably the courts should not use these cases when they engage in comparative proportionality review. In his concurring opinion in \textit{State v. Furman}, Justice Utter made reference to pre-\textit{Furman} trial judge reports filed for juvenile defendants in support of his argument that Furman’s sentence should be reversed not simply on the grounds indicated there, but also on the grounds of disproportionality. \textit{Id.} at 461–67, 858 P.2d at 1104–08 (Utter, J., concurring). Specifically, Justice Utter noted that none of the seven other defendants convicted of aggravated first-degree murder since 1981 had received the death penalty. \textit{Id.} at 462–67, 858 P.2d at 1105–07 (Utter, J., concurring). Much the same argument, incidentally, can be made regarding defendants who were convicted of aggravated first-degree murder, but have also been deemed “mentally retarded,” as defined in RCW 10.95.030(2)(a). Because the statute stipulates that no person falling into this category can be sentenced to death, as with juveniles, such defendants should be included in the Washington State Supreme Court’s database of trial judge reports because such reports must be filed for all aggravated first-degree murder convictions. However, they should not be considered when the court conducts its comparative proportionality reviews because the trial court cannot consider imposition of the death penalty in these cases. To the best of my knowledge, the only defendant who fits into this category is Mario Ortiz. \textit{See Report of the Trial Judge No. 1 (Mario Ortiz).}

\textsuperscript{226} In \textit{State v. Woods}, 143 Wash. 2d 561, 23 P.3d 1046 (2001), the Washington State Supreme Court expressly rejected the claim that the trial judge reports must be updated after any given death penalty is overturned or the case in question is remanded for resentencing by an appellate court. \textit{Id.} at 612–14, 23 P.3d at 1074–75. To do so, the court determined that the language of RCW 10.95.120 does not confine the universe of cases to be considered for the purposes of comparative proportionality review to those that are upheld on appeal. \textit{Id.} at 613, 23 P.3d at 1075. While this is technically true, the effect of the court’s statutory literalism is to permit it to cite, in its future reviews, death sentences that it once upheld but later rejected. One might argue, as the court stated in \textit{State v. Elledge}, 144 Wash. 2d 62, 26 P.3d 271 (2001), that the trial judge reports should be revised only if a sentence is rejected by an appellate court on the ground of disproportionality. \textit{Id.} at 79 n.5, 26 P.3d at 281 n.5. That will not do, however, because whatever errors occurred at the guilt and/or sentencing phase(s) of a capital trial may well have affected the jury’s ultimate decision. Because, as indicated in Part IV.B below, the court is committed to the view that it should not second-guess the jury’s decision unless there is compelling reason to do so, the court’s refusal to update the trial judge reports means that the errors remain on the record.


\textsuperscript{228} The National Center for State Courts recommends removal of such reports. \textit{User Manual, supra} note 123, at A-8.
Conduct of Comparative Proportionality Review

dead. Additionally, if the court justifies a decision to uphold a death sentence on the basis of information contained in the trial judge reports for any of these cases, its comparison will be invalid. To do so, as Justice Handler of the New Jersey State Supreme Court observed, “produces the anomaly that a reversed sentence, by definition too unreliable to carry out on the sentenced defendant, is yet reliable enough for the purpose of comparison with other capital sentences in proportionality review.”

Finally, a majority of the trial judge reports included in the supreme court database were not submitted “within thirty days after the entry of the judgment and sentence,” as required by law. Specifically, of the 259 trial judge reports, 161 (62%) were submitted late. The period of lateness ranges from a low of two days to a high of just under eight years. The latter instance may be extreme, but is by no means aberrant. A total of seventy-nine, just under 31% of all reports, were received over a year late; fifty-eight over two years; forty over three years; and twenty-six over four years. The 161 late reports were received an average of 660 days past the statutory deadline. This chronic tardiness in the filing of reports raises serious questions about the compliance of trial judges with the imperatives of the law, and, as the next subsection explains, poses equally serious questions about the accuracy of the information these reports contain.

233. It appears that the State Supreme Court sought to remedy the problem of late filings through adoption of a court rule in December 1997. See Superior Court Special Proceedings Rules—Criminal 6, available at http://courts.wa.gov/court rules (last visited July 20, 2004). This rule requires the prosecutor and defense counsel to complete and submit to the clerk of the trial court a copy of the questionnaire specified in RCW 10.95.120 within fourteen days following the entry of a judgment and sentence for an aggravated first-degree homicide conviction. Id. Once completed and served on each of the opposing attorneys, that questionnaire assists in the preparation of the trial judge’s report. That report, as the statute and rule require, must be filed within thirty days after entry of the judgment and sentence. Wash. Rev. Code § 10.95.120. Perhaps in response to the promulgation of this rule, a flurry of very late reports (a total of ten, averaging just over five years past their respective due dates) were filed with the State Supreme Court during the final month of 2001 and the first three months of 2002. On balance, however, it would appear that this rule has had little effect in expediting submission of the trial judge reports. Of the seventy-five reports filed since this rule became effective on January 1, 1998, only twenty-three have been submitted in accordance with the statutory deadline.
2. Inaccurate and/or Inadequate Information Regarding Defendants and Victims

No doubt in part because so many have been filed late, a very large number of trial judge reports provide insufficient and/or inaccurate information which cannot help but compromise the State Supreme Court’s ability to conduct meaningful comparative proportionality review. In several reports, for example, as a result of the passage of time, trial judges found themselves unable to complete many of the questions on the form. Two examples, taken from the many that might be cited, serve to illustrate this problem. First, in Report No. 210 (Cheyenne Brown), which was submitted approximately three and a half years late, the trial judge explained that the mitigation packet prepared by defendant’s counsel, containing “official records from many state agencies in Washington and Minnesota, school records, and extensive psychological records,” was no longer available to her. Therefore, she continued, “where I do not have a specific recollection of a fact I will indicate unknown rather than guess.” The judge then proceeded to enter “unknown” in response to fifteen different questions in the report. Second, in response to the report’s question regarding the number of jurors of the same race as the defendant and/or victim, the judge who completed Report No. 90 (Gerald Hankerson), concerning an African-American co-defendant sentenced to life imprisonment for murdering a Vietnamese immigrant, indicated the following: “I cannot recall if any black jurors served.” While this response is not surprising given that this form was submitted some three years after sentencing, it raises significant questions about the Washington State Supreme Court’s ability to investigate the role of racial prejudice in sentencing patterns for aggravated first-degree murders.

235. Id.
236. Id.
237. Report of the Trial Judge No. 90 (Gerald Hankerson).
238. See Report of the Trial Judge No. 95 (Kenneth Schrader), submitted more than five years late. In response to the same question about jury composition, the judge wrote: “I do not honestly recall if any jurors were members of minority races. It is possible that one or two may have been.” Id.; see also Report of the Trial Judge No. 208 (Kenneth Comeslast), submitted approximately five years late, in which the judge wrote, “I do not recall,” in response to the same question. Id. Another example is the Report of the Trial Judge No. 104 (Harold Eirich), submitted approximately three years late, in which the judge wrote, “[n]o record of race or ethnic origin of jurors has been kept.” Id. A comment of this sort obviously raises questions about whether the judge who presided over
Conduct of Comparative Proportionality Review

In many more instances, rather than confessing to a lapse of memory, judges have simply not furnished answers to questions included in the form, despite the express instruction: “Please answer each question. If you do not have sufficient information to supply an answer, please so indicate after the specific question.” Indeed, when all 259 trial judge reports are considered, the only questions that are answered without exception are those that ask for the defendant’s name, the plea he or she entered, and whether representation was provided by counsel.

Perhaps the most troublesome instance of a question left unanswered is the final question of section (6) (General Considerations). The task force that drafted the form from which the present version is derived indicated that the questionnaire’s purpose is to “aid the Supreme Court in its review of a death sentence by providing the Court with all possible information regarding the defendant and the proceedings, particularly by the trial was the same judge who completed the trial judge report. No guesswork is necessary to confirm this hypothesis in the case of Report of the Trial Judge No. 226 (Marvin Leo). There, the judge who completed the form wrote, “[t]he trial judge failed to complete this report. He was defeated in his bid for re-election and I succeeded him as Dept. 2 judge for Pierce County. They have been unable to get Judge Rudy Tollefson to do this report.”

Moreover, and contrary to the clear implication of RCW 10.95.120, which seeks “general comments of the trial judge concerning the appropriateness of the sentence,” see App. A, infra p. 880, at (6)(k), and which requires the trial judge to sign and date the questionnaire, see App. A, infra p. 880, it appears that some reports have been completed by court staff, long after the fact, rather than by the judge who presided over the trial. In one trial judge report, question marks appear throughout the report and all questions seeking information beyond that included in the record itself are simply left blank. Report of the Trial Judge No. 211 (Dennis Anfinson). In apparent explanation, the person who completed the report indicated that the “[j]udge is retired—file contained very little information.” Id. Finally, in at least one instance, the trial court judge indicated that the “[t]rial judge, after completing this report to the best of her information and belief, forwarded the report to the prosecutor for completion.” See Report of the Trial Judge No. 121 (Sterling Jarnagin). There is no indication of whether the report was also submitted to the defense counsel as required by RCW 10.95.120.

239. App. A, infra p. 868. In some instances, the absence of a response to a question is not problematic. For example, section (1)(f) of the trial judge report asks whether a psychiatric evaluation of the defendant was performed. App. A, infra p. 870, at (1)(f). Obviously, if the answer to that question is “no,” then one should not expect a trial judge to answer the following three questions, all of which are germane only if the answer to the initial question is “yes.” These questions concern the defendant’s ability to “distinguish right from wrong,” to “perceive the nature and quality of his or her act,” and to “cooperate intelligently in his or her own defense.” By the same token, in cases where the death penalty was not sought by the prosecutor, all of the questions regarding the special proceeding conducted to determine whether the sentence shall be death or life imprisonment without possibility of parole become irrelevant and so unnecessary to answer.

240. In what follows, I have elected to focus my inquiry on those questions left unanswered that are most problematic from the standpoint of comparative proportionality review. Many other examples, though, might be cited. For example, in seventeen of the reports (6%), no answer is provided in response to the request for the date on which the report itself was completed.
eliciting from the trial judge his own unique perspective of the trial.”

Despite this exhortation, in forty-seven of the 259 reports (19.1%), trial judges have failed to offer any response to the question asking for “general comments . . . concerning the appropriateness of the sentence, considering the crime, the defendant, and other relevant factors.”

Furthermore, when that question is answered, the response is often perfunctory and uninformative: “[t]he sentence was appropriate”; “[t]he legislature determined the sentence for this defendant. I do not choose to comment on the legislature’s judgment”; and “[t]he sentence was entirely appropriate given the gravity of the offense.”

Problems of the same sort are apparent in all other sections of the reports on file with the court. For example, section (1) solicits information about the defendant.

In response to the questions included in this section, Report No. 198 (Joseph Schuler II) indicates that the defendant is a male African-American, but provides virtually no other information about him. Specifically, this report fails to furnish answers to the following questions: the defendant’s date of birth and marital status; number and age of children; whether the defendant’s father and/or mother are alive or dead; the number of children born to defendant’s parents; information regarding the defendant’s formal education, intelligence level, and IQ score; pertinent psychiatric or psychological information; the work record of the defendant; prior convictions; and the defendant’s length of residence in Washington and the county of conviction. Quite remarkably, this report also provides


242. Examples include section (6)(k) in the following trial judge reports: Report of the Trial Judge No. 18 (Brian Kester); No. 19 (Michael Cornethan); No. 35 (David Lennon); No. 46 (Nicholas Giesa); No. 48 (Christopher St. Pierre).

243. Report of the Trial Judge No. 6 (Kenneth Hovland).

244. Report of the Trial Judge No. 11 (Rosaline Edmondson).

245. Report of the Trial Judge No. 40 (John Schoenhals). On rare occasions, a trial court judge will indicate that he or she disagrees with a sentence imposed by a jury. But typically, the judge’s elaboration of the reasons for the disagreement is not helpful. Consider, for example, Report of the Trial Judge No. 2 (Arnold Brown), in which the judge answered 6(k) as follows: “[s]ince there were no mitigating circumstances the failure of the jury not to impose the death penalty was not appropriate.” See also Report of the Trial Judge No. 21 (Leslie Pounds) (“Because of unusual factual situation life without parole seems a little harsh (defendant’s lack of recollection—immunity granted participant whose involvement was substantial—prior good character of defendant.”).


247. See Report of the Trial Judge No. 198 (Joseph Schuler II).

248. Id. at (1)(a)-(j).
Conduct of Comparative Proportionality Review

no answers to the questions in section (7) (Information about the Chronology of the Case), which ask for the dates of the defendant’s offense, arrest, trial, and sentence.²⁴⁹

While this report represents an extreme example of apparent judicial negligence, inspection of the remaining trial judge reports turns up many others nearly as inadequate in providing basic information about the biography and character of defendants.²⁵⁰ For example, in forty-one of the 259 reports (15.8%), no answer is provided in response to the question regarding the highest grade completed by the defendant; in sixty-three reports (24.3%), to the question regarding the defendant’s intelligence level; and in 149 reports (57.5%), to the question regarding the defendant’s IQ score. These deficiencies raise vexing questions about the capacity of the court to compare the character of various defendants, to make informed judgments about mental capacity, and so to determine the degree of culpability to be ascribed to one defendant as opposed to another.

Somewhat less common, but by no means infrequent, are those reports in which questions are left unanswered in section (4) (Information about the Victim), which is intended to provide the supreme court with a more adequate understanding of the crime in question, and, specifically, its more or less heinous nature. For example, Report No. 145 (Cristian Delbosque) fails to answer the questions that ask whether the victim was related to the defendant by blood or marriage; for the victim’s occupation as well as whether the victim was an employer or employee of the defendant;²⁵¹ whether the victim was

²⁴⁹. Id. at (7).

²⁵⁰. In section (1) of the 259 trial judge reports, date of birth is omitted in twelve (4.4%); gender in seven (2.7%); marital status in eleven (4.2%); race in three (1.2%); number and ages of defendant’s children in eight (3.1%); whether defendant’s father is living in eight (3.1%); whether defendant’s mother is living in eleven (4.2%); number of siblings of defendant in twelve (4.6%); whether a psychiatric evaluation was performed on the defendant in nine (3.5%); the existence of character or behavioral disorders on the part of the defendant in twenty-eight (10.8%); the work record of the defendant in seven (2.7%); the defendant’s prior convictions, if any, in sixteen (6.2%); and, finally, the length of time the defendant has resided in Washington in nine (3.5%). For other examples of reports that provide minimal information about the defendant, see Report of the Trial Judge No. 99 (Stanley Runion); No. 103 (José Nash); No. 124 (Christopher Bradley); No. 189 (Kevin Boot); No. 193 (David Lewis); No. 197 (Joseph Revay); No. 202 (Gary Ackley); No. 207 (Donnie Ivy); No. 211 (Dennis Anfinson); No. 212 (Sap Kray); No. 213 (Richard Morgan); No. 221 (Steven Phillips); No. 228 (Rosendo Delgado); No. 232 (Kenneth Leuluiali’i); and No. 250 (Joseph Kennedy).

²⁵¹. Report of the Trial Judge No. 145 (Cristian Delbosque) at (4)(a) & (b).
acquainted with the defendant; for the victim’s length of residency in Washington; whether the victim was the same race or ethnic origin as the defendant; whether the victim was of the same sex as the defendant; whether the victim was held hostage during the crime; whether physical harm or torture was inflicted upon the victim prior to death and, if so, of what sort and for how long; for the age of the victim; and, finally, for the type of weapon employed in the crime. Again, while this is perhaps an extreme example, many other reports provide very little information regarding the biography of the victim and the way in which he or she was murdered, once again leaving the court ill-equipped to complete its task of comparing defendants convicted of “similar” crimes. Not surprisingly, moreover, the reports that provide very little information regarding the defendant are also those that provide very little information on the nature of the crime and its victim, thus making it effectively impossible for the court to derive any benefit from these reports.

Finally, yet another variation on the problem of incomplete reports are those whose responses are almost if not entirely limited to questions that can be answered by checking a “yes” or “no” box; or, alternatively, those that furnish a bare minimum of comment, but only in response to those questions that are straightforwardly factual in nature (e.g., number and ages of defendant’s children and defendant’s prior convictions). Such reports are largely superfluous for the purposes of proportionality review because Washington’s capital punishment statute stipulates that within ten days of the entry of a judgment and sentence imposing the death penalty the clerk of the trial court must transmit to the supreme court information regarding “the caption of the case, its cause number, the defendant’s name, the crime or crimes of which the defendant was convicted, the sentence imposed, the date of entry of judgment and

252. Id. at (4)(c). In some reports, the answer provided to this question is suspect. For example, in Report No. 208 (Kenneth Comeslast), which was filed just short of five years late, the judge answered, “I believe the victims were acquainted with the defendant but at this time I cannot recall the extent of their acquaintance.” Report of the Trial Judge No. 208 (Kenneth Comeslast).

253. See Report of the Trial Judge No. 145 (Cristian Delbosque) at (4)(d)–(j).

254. For examples of additional reports that provide very little information on the nature of the crime and its victim, see Report of the Trial Judge No. 197 (Joseph Revay); No. 202 (Gary Ackley); No. 211 (Dennis Anfinson); No. 212 (Sap Kray).

255. For examples of reports that are limited to box checking and/or the provision of bare factual information, see Report of the Trial Judge No. 15 (Patrick Jeffries); No. 22 (Fortunato Dictado); No. 27 (John Anderson); No. 28 (Gus Turner).
Conduct of Comparative Proportionality Review

sentence, and addresses of the attorneys for the parties."256 Many trial judge reports, in other words, do little more than replicate the information included in the notice submitted by the trial court clerk and, as such, are of little or no use.

3. Inaccurate and/or Inadequate Information Regarding Aggravating and Mitigating Factors

The absence of adequate information in the trial judge reports concerning aggravating and mitigating factors particularly compromises the ability of the Washington State Supreme Court to conduct meaningful comparative proportionality review. Some trial judge reports fail to indicate the statutorily specified circumstances that warranted conviction on a charge of aggravated first-degree murder. A considerably larger number fail to provide adequate information regarding any mitigating evidence that may have been introduced during special sentencing proceedings.

At least one applicable aggravating circumstance must be found in order to sustain a conviction for aggravated first-degree murder, regardless of whether that conviction generates a sentence of death or life without parole, and regardless of whether a conviction results from a guilty plea or a jury determination.257 The inclusion of twelve reports (4.6% of the total) that do not indicate any aggravating circumstances, alleged or found applicable, indicates either that these reports do not belong in the database maintained by the court or, that they incorporate a serious legal error.258 This omission makes it impossible to know the

256. WASH. REV. CODE § 10.95.100 (2000).
257. Id. § 10.95.020. Some trial court judges seem to believe that a guilty plea exempts them from indicating the aggravating circumstance(s) that provides the basis for conviction on a charge of aggravated first-degree murder. Because the question concerning aggravating circumstances appears in the section of the trial judge reports regarding the trial (as opposed to the section on the special sentencing proceeding), this belief is without foundation. In two cases, in response to the questions regarding alleged and applicable aggravating circumstances, the judge wrote “Not applicable,” and explained this entry by referring back to the question concerning the defendant’s initial plea. Report of the Trial Judge No. 152 (Steven McCord); No. 153 (Ernest Benson). Moreover, in another report, a judge listed “more than one victim and murders part of a common scheme” under the heading of alleged aggravating circumstances, but then checked two boxes under the heading “Found Applicable.” Report of the Trial Judge No. 196 (Barry Loukaitis). In fact, according to Washington State’s death penalty statute, this is a single aggravating circumstance. WASH. REV. CODE § 10.95.020(10). This error effectively makes this homicide appear more heinous when measured by number of aggravating circumstances than it was in fact.
258. For trial judge reports that fail to indicate any aggravating circumstances, alleged and/or applicable, see Report of the Trial Judge No. 24 (Grady Mitchell); No. 55 (Russell Stenger); No. 80
basis on which a defendant was convicted of aggravated first-degree murder and thereby deprives the Washington State Supreme Court of one of the principal means of assessing the heinousness of any given murder.

The problem posed by inadequate information regarding mitigating circumstances is far more consequential because the supreme court, in reviewing a death sentence, is statutorily required to assess the evidence that persuaded a jury not to afford leniency to a defendant. At the close of a special sentencing proceeding, jurors must answer the following question: “Having in mind the crime of which the defendant has been found guilty, are you convinced beyond a reasonable doubt that there are not sufficient mitigating circumstances to merit leniency?”

Because the jury’s attention is directed by this question to the presence or absence of mitigating circumstances as the primary factor to consider in its deliberations, the failure of the vast majority of trial judge reports to elicit any significant detail on this issue seriously compromises the meaningfulness of the supreme court’s conduct of comparative proportionality review.

Two questions in the trial judge report deal with mitigating circumstances introduced during the penalty phase of capital trials. The first asks whether there was, “in the court’s opinion, credible evidence of

260. Id. § 10.95.060(4) (2000). Incidentally, one report includes what is arguably a valid complaint regarding the formulation of this statutory question: “The question asked of the jury . . . is asked in the negative, and the jurors had difficulty understanding the concept, as well as the question, and the application thereof. It is confusing!” Report of the Trial Judge No. 77 (Charles Tate).
261. The absence of adequate evidence regarding mitigating circumstances is also problematic if a prosecutor elects to consult the trial judge reports for assistance in determining whether or not to seek the death penalty. RCW 10.95.040 specifies that “[i]f a person is charged with aggravated first degree murder as defined by RCW 10.95.020, the prosecuting attorney shall file written notice of a special sentencing proceeding to determine whether or not the death penalty should be imposed when there is reason to believe that there are not sufficient mitigating circumstances to merit leniency.” WASH. REV. CODE § 10.95.040 (2000). Should a prosecutor seek guidance from the trial judge reports in deciding whether or not to seek the death penalty, little will be gleaned from this source.
Conduct of Comparative Proportionality Review

any mitigating circumstances as provided in Laws of 1981, ch. 138, § 7.262 The second asks whether there was “evidence of mitigating circumstances, whether or not of a type listed in Laws of 1981, ch. 138, § 7, not described” in the answer to the previous question.263 The first of these questions is intended to elicit information about mitigating factors enumerated by the state legislature (e.g., “[w]hether the defendant acted under duress or domination of another person;” “[w]hether, at the time of the murder, the capacity of the defendant to appreciate the wrongfulness of his or her conduct . . . was substantially impaired as a result of mental disease or defect;” and/or “[w]hether there is a likelihood that the defendant will pose a danger to others in the future”).264 The second of these questions is included in order to comply with the U.S. Supreme Court’s ruling in Lockett v. Ohio.265 That ruling, requiring individualized sentencing in capital cases, permits a defendant to offer as evidence during the sentencing phase “any aspect of [his or her] character or record and any of the circumstances of the offense” that may warrant a sentence other than death.266

264. WASH. REV. CODE § 10.95.070(5), (6), and (8) (2000).
266. Id. at 604. Because both of these questions appear in section (3) of the trial judge questionnaire (Information Concerning the Special Sentencing Proceeding), App. A, infra pp. 873–74, at (3), information regarding mitigating circumstances is only solicited from judges when the death penalty is sought by the state and a special sentencing proceeding is in fact conducted. So constructed, it is not clear that the two questions in section (3) adequately satisfy the principal purpose of the trial judge reports. If that purpose is to provide the State Supreme Court with as much information as possible regarding the factors that justified a sentence of death as opposed to life imprisonment without the possibility of parole, and if that information is a necessary condition of meaningful comparative proportionality review, then it is at least arguable that the form should also seek to draw out any information relevant to this question that was introduced during the guilt phase. Yet, in section (2) (Information about the Trial), trial judges are simply given a list of possible defenses (e.g., excusable homicide, insanity, duress, alibi, intoxication), followed by one column of boxes in which they must indicate whether evidence was introduced regarding any of these defenses. A second column follows in which they must indicate whether the jury received instructions regarding these defenses. App. A, infra pp. 871–73, at (2). This checklist fails to solicit any information introduced during the guilt phase that may bear on the question of mitigating circumstances, should a special sentencing proceeding ensue. For example, should a defendant introduce evidence regarding the defense of duress, that information will almost certainly prove pertinent during the special sentencing proceeding to the statutorily enumerated mitigating circumstance dealing with the question of whether a defendant “acted under duress or domination of another person.” WASH. REV. CODE § 10.95.070(5). Yet the trial judge report does nothing to secure information about this evidence, unless of course it is reintroduced during the special sentencing proceeding and the trial court judge, in completing the report, sees fit to introduce that evidence in response to the questions concerning that proceeding.
Of the cases in which the death penalty was sought, and a special sentencing proceeding was conducted, only very rarely do the trial judge reports indicate in any detail just what mitigating evidence was introduced by the defense. In eight instances, one or both of the questions concerning mitigating evidence has been left entirely unanswered (9.7% of the total), leaving the Washington State Supreme Court uninformed about a key component of special sentencing proceedings. In an additional eleven instances (13.4%), the “No” box is checked in response to both of these two questions. In such instances, leaving aside the very small number of death penalty “volunteers,” it is difficult to know whether these negatives are indicative of a failure on the part of defense counsel to introduce such evidence, which raises questions about the adequacy of legal representation in death penalty cases, or whether they are indicative of a failure on the part of the trial judge to record that evidence.

More significantly, in over half of these reports, when a judge has indicated that credible mitigating evidence has been introduced, his or her account of that evidence is limited to a single phrase or sentence. To provide just three examples, in Report No. 174 (Timothy Blackwell), the trial judge wrote: “The life history of the defendant;” in Report No. 92 (Marc Darrah), the judge indicated only that “the defendant & prosecutor stipulated that there was [sic] mitigating circumstances”; and in Report No. 164 (Steven Morgan), the judge reported that the “defendant’s family and several ministers testified,” but gave no indication of the nature of that testimony. Moreover, although the form expressly asks for evidence regarding mitigating circumstances that the court found “credible,” in several reports, trial judges have cited such evidence only to discount its significance. Thus, in Report No. 140 (Cal Brown), the judge did not indicate in any substantive way the mitigating evidence that was introduced, limiting his assessment to the

267. In only three instances would I describe the trial judge’s account of mitigating circumstances as adequate. See Report of the Trial Judge No. 70 (Herbert Rice); No. 73 (Michael Furman); No. 220 (Dayva Cross).
269. Report of the Trial Judge No. 92 (Marc Darrah).
270. Report of the Trial Judge No. 164 (Steven Morgan). For other examples of trial judge reports which provided extremely limited account of mitigating circumstances, see Report of the Trial Judge No. 45 (James Dykgraaf); No. 75 (Gary Benn); No. 140 (Cal Brown); No. 175 (Richard Clark); No. 224 (Nicolas Vasquez).
Conduct of Comparative Proportionality Review

following after checking the “yes” box: “but . . . it did not appear to be sufficient to merit leniency.”

Similarly, in Report No. 181 (Henry Marshall), the judge noted: “There was evidence of a mitigating factor set out in the law, but given the other facts of the case it was not particularly credible as it all related to his ability to plan and carry out a plan. The planning of the robbery and the carrying out of the robbery indicated that he performed these functions quite well.”

When this difficulty is considered in combination with the impoverished quality of the evidence regarding mitigating circumstances in the trial judge reports, it seems clear that the Washington State Supreme Court is in no position to reliably determine why any specific case resulted in a sentence of life imprisonment rather than death.

4. Inaccurate and/or Inadequate Information Regarding Race and Ethnicity

The trial judge report form does not provide a standardized set of categories for determining the racial and ethnic identities of various participants in aggravated first-degree murder trials. Also judges do not have any reliable means of determining those identities. As a result, in conducting its death sentence reviews, the Washington State Supreme Court is unable to gauge the presence or absence of discrimination in the administration of the death penalty in any systematic way and is thus unable to meet one of the two primary objectives of comparative proportionality review. This is true with regard to efforts to compare various cases among the trial judge reports, and it is equally true with


273. Report of the Trial Judge No. 181 (Henry Marshall). Ironically, Henry Marshall’s death sentence was subsequently vacated by the State Supreme Court when it held that the trial court erred in refusing to allow the defendant to withdraw a guilty plea when presented with significant evidence of his mental incompetence. State v. Marshall, 144 Wash. 2d 266, 277–81, 27 P.3d 192, 198–200 (2001).

274. I am fully aware of the problematic nature of such racial and ethnic classifications. My point is not to commend their employment, but, rather, to indicate how their absence from the trial judge reports undermines the effort to ferret out instances when constitutionally impermissible factors influence the deliberation and judgment of juries. For specification of the standardized categories employed by the federal government, see U.S. Dep’t of Commerce, Bureau of the Census, United States Census 2000, available at http://www.census.gov/dmd/www/pdf/d61a.pdf (last visited June 29, 2004).

275. State v. Lord, 117 Wash. 2d 829, 910, 822 P.2d 177, 223 (1991) (“Our concern is with alleviating the types of major systemic problems identified in Furman: random arbitrariness and imposition of the death sentence based on race.”).
In regard to efforts, internal to a single case, to determine the import of responses to questions that ask whether the victim, jurors and/or witnesses are or are not of the same race or ethnic identity as the defendant.

Section (6) of the trial judge report requests information regarding the race or ethnic origin of the defendant, the victim, the jury, and the population of the county in which the trial was conducted. Related questions ask whether any evidence indicates that persons of a particular race or ethnic origin were systematically excluded from the jury; whether the race or ethnic origin of the defendant, victim, or any witness was an apparent factor at trial; and whether the jury was expressly instructed to exclude race and ethnic origin in its deliberations. These questions are intended to elicit the sort of information required by the court if it is to respond to the concern, articulated in *Furman*, that impermissible forms of bias may explain why some are sentenced to die, whereas most convicted of aggravated first-degree murder are not.

The responses in the trial judge reports to the questions regarding racial and ethnic identity include, but are not limited to, the following: African; African-American; Black; Black (Father Black, mother Caucasian); Albanian Muslim; American Indian-Colville; Anglo-Saxon; Asian; Asian (Filipino); Cambodian; Caucasian; Caucasian/Cajun; Caucasian-North American; European/N. American; Hispanic; Indian; Latino (Mexican); Mexican-American; Native; Native American; Samoan; Thai; Unknown; White; and White American. Setting aside the racism and sexism evident in certain of these categorizations (e.g., Black-Father Black, mother Caucasian); and disregarding the failure in

278. See Report of the Trial Judge No. 1 (Mario Ortiz) (Mexican American); No. 2 (Arnold Roy Brown) (Caucasian); No. 3 (Dwayne Earl Bartholomew) (White); No. 12 (Gregory Tyrone Brown) (Black); No. 13 (Kwan Fai Mak) (Asian); No. 42 (Kenneth R. Petersen) (European/N. American); No. 54 (Sandra Entz) (Anglo-Saxon (Caucasian)); No. 70 (Herbert A. Rice, Jr.) (American Indian-Colville); No. 71 (Patrick G. Hoffman) (Indian); No. 73 (Michael Monroe Furman) (Caucasian-North American); No. 77 (Charles Curtis Tate) (Black (Father Black, Mother Caucasian)); No. 79 (Sherwood Knight) (African American); No. 80 (Gabriel Garcia) (Hispanic); No. 114 (Kly Bun Meas) (Cambodian); No. 127 (Joseph Casbon) (Caucasian/Cajun); No. 130 (Cherno Camara) (African); No. 146 (Martin Rasmul Bulichi) (Albanian Muslim); No. 158 (Roderick Shawn Selwyn) (Native); No. 161 (Nga Ngoeung) (Born in Thailand); No. 197 (Joseph Revay) (Native American); No. 203 (Marvin Jay Francisco) (Asian (Filipino)/Caucasian); No. 210 (Cheyenne Troy Brown) (Unknown); No. 224 (Nicolas Solorio Vasquez) (Latino (Mexican)); No. 232 (Kenneth John Leuluiaialii) (Samoan); No. 235 (Aaron Eugene Howerton) (White American).
Conduct of Comparative Proportionality Review

many cases to distinguish racial from ethnic identity, as well as political or national identity from either; and, finally, leaving aside the fact that visual observation appears to be the principal means of determining to which category any given defendant, victim, jury member, or witness is assigned, this miscellany cannot help but thwart the court’s effort to determine whether Justice Douglas was correct in Furman when he stated that the death penalty is selectively applied to those who lack “political clout” and/or are members of “suspect” or “unpopular” minorities.

That trial court judges have struggled with these issues is apparent. For example, in Report No. 10 (Steven Carey), the judge identified Steven Carey as Caucasian, and, in considering the composition of the jury, he indicated that all twelve of its members were Caucasian as well. His uncertainty about this conclusion, however, is indicated by a marginal comment, which reads: “Based on visual observation; no inquiry has been made as to the race or ethnic background of the jury.” In Report No. 33 (Donald Galbert), the judge indicated that the defendant is “black”; but, then, in response to the question concerning the demographics of the county in which the trial transpired, apparently could not decide whether “black” is a racial or an ethnic category and so checked boxes indicating that it is both. Finally, in Report No. 34 (Paul Joseph St. Pierre), the judge indicated that, based on hearsay, he suspected that the defendant was Italian; but then, in response to the

279. To cite but one illustration of the latter problem, consider Report of the Trial Judge No. 214 (Gary Packer). There, the trial judge indicates that 75–90% of the population of Clark County is of the same race as the defendant (Caucasian), but that only 25–50% of its population is of the same ethnic origin as the defendant. Absent some indication of the ethnicity of the defendant, which is not provided, it is impossible to understand the distinction drawn here and so to draw any clear conclusions about the operation of prejudice in the trial and/or jury deliberation.

280. On occasion, a defendant’s, victim’s, or juror’s surname is used as a surrogate for visible appearance. See, e.g., Report of the Trial Judge No. 228 (Rosendo Delgado) (noting the trial judge’s statement that “[t]he victims and defendant (all Hispanic). It appeared there were no Hispanics on the jury, at least by surname”); Report of the Trial Judge No. 229 (Julio Delgado) (noting the same occurrence with respect to the classification “Filipino”).


282. Report of the Trial Judge No. 10 (Steven Carey).

283. Id.; see also Report of the Trial Judge No. 26 (Lonnie Link) (indicating that the defendant “appears to be Caucasian”); No. 216 (Allen Gregory) (noting the judge’s statement that “[n]o juror asked to provide his/her race or ethnic origin. Above numbers are based on appearance only.”).

284. See Report of the Trial Judge No. 33 (Donald Galbert); see also Report of the Trial Judge No. 4 (Jimmy Ramil) (noting the same occurrence with respect to the classification “Filipino”).
question concerning the composition of the jury, confessed that “I have made no effort to determine percentage of Italians.” The point to be drawn from these examples is not so much that the claims made in these reports are patently wrong. Rather, the point is that in the absence of standardized categories these claims are often arbitrary if not capricious, and, as such, cannot help but sabotage the aim of the report’s questions concerning the racial and ethnic identities of various parties to any given trial.

In addition, the absence of pre-designated categories vitiates any effort to grasp the more subtle dimensions of discrimination, i.e., those that, in principle, might be identified by examining any given sentence in light of the percentage of the county population that is the same race or ethnic identity as the defendant. In many instances, it appears that trial court judges have simply guessed when answering this question. For example, Report No. 79 (Sherwood Knight), dealing with a trial conducted in 1986, indicates that African-Americans in King County comprise 10–25% of the total population, whereas Report No. 88 (Ray Lewis), concerning a trial conducted in the same year, indicates that the African-American population of King County is under 10%. Both claims, to belabor the obvious, cannot simultaneously be true. Similarly, Report No. 109 (Gregory Scott), filed in 1993, states that the Caucasian population of King County makes up 50–75% of the total population, whereas the 1990 census indicates that the correct figure is just shy of 85%. Additionally, Report No. 193 (David Lewis) states that the African-American population of Pierce County in 1999 is between 10% and 25%, whereas the 2000 census indicates that the total

285. Report of the Trial Judge No. 34 (Paul Joseph St. Pierre). In some cases, the answers provided to these questions are simply confusing. In one report, the defendant is identified as Caucasian. See Report of the Trial Judge No. 4 (Dennis Williams). Then, in response to the question concerning the demographics of the county in which the trial was held, the judge indicated that over 90% of the population is of the same race or ethnic origin as the defendant, and, in apparent explanation for this claim, contends that “there is a significant Native American population in Mason County.” Id.

286. See Report of the Trial Judge No. 79 (Sherwood Knight); No. 88 (Ray Lewis).

287. In this instance, the second judge got it right. The 1990 U.S. Census data indicates that African-Americans made up 5% of the total population of King County. See 1990 US Census Data, at http://homer.ssd.census.gov/cdrom/lookup/1086358727=/1086358739=/1086358746=/1086358787=/CMD=RET/DB=C90TF3A/F0=FIPS.STATE/F1=FIPS.COUNTY90/F2=STUB.GEO/FMT=HTML/LEV=COUNTY90/SEL=53,033,King+County/T=P8 (last visited June 29, 2004).

288. Report of the Trial Judge No. 109 (Gregory Scott); see 1990 US Census Data, supra note 287.
Conduct of Comparative Proportionality Review

was 7%. In such instances, the fault lies not so much with the trial court judges but with the form itself and, more particularly, its presupposition that judges are in a position to know and to certify the accuracy of demographic information that is not within their professional ken.

To the extent that the trial judge reports are defective in the ways indicated here, the Washington State Supreme Court’s ability to conduct meaningful comparative proportionality reviews is compromised. Moreover, even if the supreme court were to attempt to remedy the deficiencies in the database, it could not do so. The accuracy and utility of these reports depend on the trial judges’ conscientious completion and timely submission. However, as demonstrated above, this condition has not been met in a majority of the reports, with the possible exception of those relating to cases most recently decided. One could attempt to remedy the failings of the court’s database by returning to the trial record of every past aggravated murder case in order to seek answers to questions that are now unanswered, incomplete, or erroneous. That, however, would be a largely pointless gesture because the principal purpose of the trial judge reports is to collect information that goes beyond what is contained in those records.

B. Comparative Proportionality Review in Washington and Its Defects

Since adoption of the current death penalty statute in 1981, the Washington State Supreme Court has struggled to develop a coherent and consistent method to inform its conduct of comparative proportionality review. It has largely failed in this effort. As a result, its performance of such reviews has remained an ad hoc affair that does

289. Report of the Trial Judge No. 193 (David Lewis); 2000 US Census Data, at http://quickfacts.census.gov/qfd/states/53/53053.html (last visited July 3, 2004). The question of whether judges sometimes offer ill or uninformed guesses in answering this demographic query is plainly answered by Report of the Trial Judge No. 165 (Clark Elmore). There, the judge indicated that the Caucasian population of the county is 75–90% of the total population, but then proceeded to indicate that it may in fact be over 90%, adding a parenthetical comment of “possibly” in order to explain these inconsistent responses. Report of the Trial Judge No. 165 (Clark Elmore). Another judge similarly failed to answer the county demographic question, but indicated that a “big majority of Spokane [is] believed to be of white race.” Report of the Trial Judge No. 94 (Daniel Edwards); see also Report of the Trial Judge No. 235 (Aaron Howerton) (indicating that “most” of the jurors were of the same race as the defendant).

little to remedy the concerns articulated in *Furman* regarding the arbitrary and capricious administration of capital punishment. More specifically, the court has failed to develop a procedure that ensures that the death penalty is consistently applied to those convicted of the most heinous murders and, correlative, that those convicted of less heinous crimes are not sentenced to die. For the most part, moreover, the court has not engaged in the sort of inquiry necessary to ensure that constitutionally impermissible considerations such as race do not taint the administration of the death penalty. In sum, analysis of the Washington State Supreme Court’s conduct of comparative proportionality review since 1981 provides little reason to believe that this procedure has remedied the concerns that initially prompted its adoption.

During the period covered by this study, the death penalty was sought in seventy-eight cases in the State of Washington and imposed in thirty-one. Twenty sentences were reviewed for proportionality by the Washington State Supreme Court, but none were vacated for disproportionality. This section examines the defects in the court’s conduct of comparative proportionality review in those twenty cases. Prior to 1995, the court demonstrated considerable uncertainty regarding

---

291. *Furman v. Georgia*, 408 U.S. 238, 309–10 (1972) (Stewart, J., concurring); *id.* at 313 (White, J., concurring); *id.* at 253–57 (Douglas, J., concurring).


293. MARK LARRANAGA, *WASHINGTON DEATH PENALTY ASSISTANCE CENTER, WHERE ARE WE HEADING? CURRENT TRENDS OF WASHINGTON’S DEATH PENALTY* 14 n.63 (2003). Of these thirty-one, Michael Furman was a juvenile at the time of his offense and his death sentence was reversed on that basis. *id.*; see *State v. Furman*, 122 Wash. 2d 440, 458; 858 P.2d 1092, 1103 (1993).

294. The following have been analyzed for proportionality by the Washington State Supreme Court: Dwayne Bartholomew (No. 3), Charles Campbell (No. 9), Kwan Mak (No. 13), Patrick Jeffries (No. 15), Benjamin Harris (No. 29), Mitchell Rupe (Nos. 7 and 31), David Rice (No. 43), Brian Lord (No. 49), Gary Benn (No. 75), Westley Dodd (No. 76), Jonathan Gentry (No. 119), James Brett (No. 125), Blake Pirtle (No. 132), Cal Brown (No. 140), Darold Stenson (No. 144), Jeremy Sagastegui (No. 160), Clark Elmore (No. 165), Dwayne Woods (No. 177), Cecil Davis (No. 180), James Elledge (No. 183); while the following have not been analyzed for proportionality: Clark Hazen (No. 39, committed suicide while case on appeal), Michael Furman (No. 73), Sammie Luvene (No. 135), Charles Finch (No. 154), Richard Clark (No. 175), Michael Roberts (No. 176), Henry Marshall (No. 181), Covell Thomas (No. 194), Allen Gregory (No. 216), Dayva Cross (No. 220), and Robert Yates (No. 251). As of this writing, four of these defendants have been executed (Dodd, Campbell, Sagastegui, and Elledge), and three of these cases are currently on direct appeal (Gregory, Cross, and Yates). See *OFFICE OF THE ATTORNEY GENERAL, CRIMINAL JUSTICE DIVISION, CAPITAL PUNISHMENT CASE STATUS REPORT*, available at http://www.atg.wa.gov/pubs/capital_litigation/Capital_Lit_Report0304.doc (last visited July 13, 2004).
Conduct of Comparative Proportionality Review

Each of the three basic elements in the logic of comparative proportionality review: identification of the relevant universe of cases to be considered, specification of the criteria to be employed in defining "similar" cases; and, finally, articulation of a test to be used in determining whether a sentence under review is or is not disproportionate. By the close of 1995, the court concluded that the universe of cases must include all cases in which defendants were convicted of aggravated first-degree murder, and it had specified the factors it would consider in identifying a pool of similar cases. Yet, in reviewing specific cases, the court’s method of selecting pools of similar cases from this universe proved inconsistent at best, and its method of comparing those cases to those under review in order to determine proportionality proved altogether unsystematic at best and arbitrary at worst.

1. Comparative Proportionality Review to 1995

The Washington State Supreme Court’s early difficulties with all three elements of the logic of comparative proportionality review and its confusion regarding the very purpose of this practice were apparent in State v. Campbell, the first case it reviewed. In affirming the sentence and conviction of Charles Campbell, the court announced that the purpose of comparative proportionality review is to “assure that ‘wholly arbitrary, capricious, or freakish sentences’ are minimized.” In other words, rather than affirmatively seeking to ensure consistency in capital sentencing throughout the state, the court would merely seek to invalidate those deemed aberrant.

Although the court acknowledged the statutory provision requiring

296. Gilbert, supra note 217, at 604.
297. Campbell, 103 Wash. 2d at 30 n.2, 691 P.2d at 945 n.2 (quoting Pulley v. Harris, 465 U.S. 37, 45 (1984)).
298. This holding is not consistent with the legislative history of RCW 10.95.130, which indicates that the purpose of comparative proportionality review is to ensure statewide consistency in the application of the death penalty. See Memorandum from Bill Gales, Senior Counsel, Washington State Senate Judiciary Committee, to Gene Baxstrom 2 (Oct. 24, 1980). Gales stated:

The supreme court is required to review any death penalty sentence both for trial errors and for a determination of whether the sentence is excessive or disproportionate to the penalty imposed in similar cases taking into consideration both the crime and the defendant. This last requirement is intended to eliminate any possibility of differing standards for the imposition of the death penalty across the state.

Id.; see Death Penalty—Aggravated Murder, ch. 206, sec. 7(3)(b), 1977 Wash. Laws 778.
consideration of all “similar cases,” it found no trial judge report with facts comparable to those of Campbell: “While many other defendants have been charged with aggravated first-degree murder, we find no other where four aggravating factors in the guilt phase were found to be present by the jury.” Stymied in its effort to abstract a pool of similar cases from the universe defined by statute, based on the number of aggravating factors involved, the court abandoned all pretense of comparative review. Instead, it reverted to inherent proportionality review: “A case which involves such a multitude of aggravating factors, we are convinced, would, with great frequency prompt a jury to impose the death penalty . . . . Moreover, we are hard pressed to find killings more premeditated and revengeful than those committed by defendant.” Without specifying which cases and trial judge reports it had inspected, without indicating why it considered the number of aggravating factors to be the dispositive element in its review, and without explaining why it did not see any need to consider the character of the defendant, as the law requires, the court simply asserted that Campbell’s sentence was “clearly” proportionate “to the crime committed.”

In several cases following on the heels of Campbell, the court acknowledged the distinction between inherent and comparative proportionality review, as well as its statutory obligation to perform the latter. The court, however, failed to remedy the defects of its analysis in Campbell, and, in certain respects, aggravated them. For example, in the second case it reviewed for proportionality, State v. Jeffries, the court

299. Campbell, 103 Wash. 2d at 30, 691 P.2d at 945. By the end of 1983, eighteen trial judge reports had been filed with the Washington State Supreme Court, and so, at minimum, this number was available to the court when it reviewed Campbell’s case. In addition, RCW 10.95.130(b) defines as “similar” all those “reported in the Washington Reports or Washington Appellate Reports since January 1, 1965.” Whether this latter groups of cases should be considered in conducting comparative proportionality review is debatable. See supra note 217.

300. Campbell, 103 Wash. 2d at 30, 691 P.2d at 945–46.

301. Id. Upon reviewing the factual circumstances of a number of cases tried since adoption of the 1981 capital punishment statute, Justice Utter stated that the “pattern of filings and of jury verdicts under our new statute defies any rational explanation . . . . We can find no basis on the face of these cases to explain why many of the cases where the death penalty was not sought differ in substantial degree from any where the death penalty was sought.” Id. at 45, 691 P.2d at 954 (Utter, J., concurring in part, dissenting in part).

Conduct of Comparative Proportionality Review

ignored the plain language of the statute when it declared that the universe of cases to be considered does “not include cases where the death penalty was not sought by the prosecutor.” As noted above, by limiting itself to cases in which the death penalty was sought, the court neglected the far greater number of aggravated first-degree murder cases in which it is not, and so defeated the very purpose of comparative proportionality review.

The Jeffries court’s determination of the appropriate pool of similar cases was equally flawed. The court first noted the number of aggravating factors found applicable in Jeffries (two), and, second, offered summary recapitulations of four other cases in which death sentences had been imposed (absent any express reference to the trial judge reports in these cases). On that basis, absent any elaboration, the court asserted that “these four cases strongly establish that the death penalty here is not disproportionate.” No effort was made to explain why these four cases, with very different fact configurations, were deemed similar to Jeffries; why other aggravated first-degree murder cases were not deemed similar; why the number of aggravating circumstances should preclude consideration of the qualitative features of the crime and the character of the defendant; or, what criteria, if any, were employed in determining that the sentence imposed on Jeffries was indeed proportionate to these other four. Given the poverty of this review, it is difficult to escape the conclusion that in this instance the court engaged in a results-oriented inquiry that provided a veneer of statutory compliance to a sentence deemed appropriate on grounds never expressly stated in the opinion.

303. Jeffries, 105 Wash. 2d at 430, 717 P.2d at 740. In his dissenting opinion in State v. Rupe, 108 Wash. 2d 734, 743 P.2d 210 (1987), Chief Justice Pearson explained why the State Supreme Court typically confines its proportionality inquiries to cases in which the death penalty has been imposed: “In examining the special circumstances in death penalty cases, we have a substantial record before us, but in comparing death penalty cases to those in which the death penalty was not imposed we are limited to the facts elicited from the trial courts by this court’s standard questionnaire.” Id. at 786, 743 P.2d at 239 (Pearson, C.J., dissenting) (emphasis in original). Given the quality of information elicited by those questionnaires, it is perhaps not surprising (although it remains legally indefensible) that the court most often limits its comparative inquiries to death sentence cases.

304. Comparative proportionality review is meaningful only if the universe of reports includes those defendants who were convicted of aggravated first-degree murder, but for whom the death penalty was not sought, as well as those for whom the death penalty was sought but not imposed. See supra Part II.A.

305. Jeffries, 105 Wash. 2d at 430, 717 P.2d at 740.
In *State v. Harris*,306 also decided in 1986, the court tacitly acknowledged that it had erred in *Jeffries* by confining its review of cases to those in which the death sentence had in fact been imposed.307 However, as in *Campbell*, the court conceded that it could identify no other cases involving contract killings in which the death penalty had been sought: “[t]herefore, there is no evidence to be considered whether the present case is disproportionate.”308 The court, though, was able to identify three post-1981 cases involving contract killings as an aggravating factor, but in which the death penalty had not been sought.309 On the face of it, this alone would appear to argue against the proportionality of Harris’s sentence,310 especially given the court’s declaration that “it is our duty under the similarity standard to assure that no death sentence is affirmed unless in similar cases throughout the state the death penalty has been imposed generally.”311 Yet the court did not draw that conclusion. Instead, after noting that the statute “provides little guidance to determine at what point a sentence becomes proportionate or disproportionate,”312 it stated: “We are satisfied the imposition of the death penalty was not ‘wantonly or freakishly’ imposed.”313 In defense of this ruling, the court noted: “No errors were made in the guilt and sentencing phases of the defendant’s trial; the trial court made every effort to ensure the defendant’s trial proceedings were fair and just.”314 However, to offer this as a justification for affirming a death sentence is

306. 106 Wash. 2d 784, 725 P.2d 975 (1986).
307. See id. at 798, 725 P.2d at 982.
308. Id., 725 P.2d at 982–83.
309. Id. at 798–99, 725 P.2d at 983.
310. Granted, the court did distinguish *Harris* from the three other cases involving contract killings by noting that Harris had received a one-year sentence for a prior manslaughter charge, whereas the other defendants “showed little, if any, prior records that rose to the seriousness of defendant’s.” Id. at 799, 725 P.2d at 983. However, as Justice Utter noted in his dissent, two of these three murders were far more brutal than that committed by Harris. Id. at 803–06, 725 P.2d at 985–86 (Utter, J., dissenting). Moreover, in addition to disregarding the fifteen years between Harris’s prior and present convictions, the court also ignored the fact that two of the three defendants it cited had been imprisoned for various offenses much closer to the time of their convictions for aggravated first-degree murder. Id.
311. Id. at 798, 725 P.2d at 982 (quoting Moore v. State, 213 S.E.2d 829, 864 (Ga. 1975)) (emphasis in original).
312. Id., 725 P.2d at 982.
313. Id. at 799, 725 P.2d at 983.
314. Id. The court in fact erred on this point. Harris’s conviction was eventually overturned by a federal district court and he was released from prison. Harris v. Blodgett, 853 F. Supp. 1239, 1294, 1300 (W.D. Wash. 1994), aff’d, 64 F.3d 1432 (9th Cir. 1995).
Conduct of Comparative Proportionality Review

to abdicate the court’s statutory mandate to conduct an independent assessment of proportionality. As Bruce Gilbert noted:

The jury has never made a factual determination on the proportionality of the sentence. They have looked at no similar cases, and are presumably basing their decision solely on the facts of their individual case. While the presumption that a jury sentence is valid is relevant to whether the defendant deserves to die, or any other factual determination the jury has been asked to decide, comparative proportionality review is not such a determination.315

In State v. Rupe,316 decided in 1987, the court finally conceded that the Washington statute requires it to consider all convictions for aggravated first-degree murder, regardless of whether the death penalty was sought or imposed.317 Resolution of this first issue in the logic of comparative proportionality review did not, however, lead to successful resolution of the second (determination of the pool of cases to be deemed “similar” to any specific case under review), or, for that matter, the third (determination of what conditions must be met in order to find a sentence disproportionate). With respect to the second, the Rupe court could find only a single death sentence case (David Rice), which, although not yet reviewed by the court, exhibited the same combination of aggravating factors (multiple victims were murdered as part of a common scheme or plan; the murders were committed in order to conceal the perpetrator’s identity; and, lastly, they were committed in the course of a first-degree robbery).318 Apparently unwilling to rest its conclusion on a single case it had yet to review, again without explanation, the court declared that the presence of two of the three aggravating factors found applicable in Rupe sufficed to render other cases comparable.319 Employing this criterion, the court then identified four cases in which defendants had received the death sentence; two in which defendants were sentenced to life without parole, although the prosecution had sought the death penalty; one in which the prosecutor did not seek the death penalty; and another in which the defendant

317. See id. at 767–68, 743 P.2d at 229.
318. See id. at 768, 743 P.2d at 229.
319. See id. at 768–69, 743 P.2d at 229–30.
pleaded guilty and was sentenced to life without parole. Without explaining why an even split in these sentencing decisions was sufficient to satisfy the criterion of “generality,” the court affirmed Rupe’s sentence. Rupe’s sentence was eventually reversed, leading to his re-sentencing to life without parole.

In addition, the Rupe court conceded that RCW section 10.95 requires it to consider the character of the defendant as well as that of his or her crime in deciding questions of proportionality. Reading Rupe’s character as an additional justification for a death sentence proved to be a tricky matter, however, because, as the court acknowledged, his “background is unusual among those convicted of first-degree aggravated murder”: “Rupe has no prior criminal record. There is evidence that, most of his life, he was active in his community, respected and liked by others.” However, the court discounted this mitigating evidence, stating that “[t]he Legislature has clearly contemplated that the death sentence is appropriate for crimes such as he has committed.”

Deference to the legislature, however, is no more defensible than is deference to a jury. In both instances, the court fails to engage in an independent inquiry regarding whether any given death sentence constitutes a departure from the sentences generally imposed on other defendants, considering both the crime and the defendant.

The court offered yet another twist in its ongoing struggle to determine the relevant criteria of similarity when it affirmed the death

---

320. See id.
321. Id.
322. State v. Rupe, 93 F.3d 1434, 1437 (9th Cir. 1996).
324. See Rupe, 108 Wash. 2d at 770, 743 P.2d at 230.
325. Id. During Rupe’s sentencing hearing, approximately fifty persons testified on his behalf. Id. at 780, 743 P.2d at 235. He had been involved in many activities, including the Boy Scouts, the Civil Air Patrol, and the Mason County Search and Rescue Council. Id. at 780, 743 P.2d at 235–36. Moreover, he had served in the Army, had never previously committed a crime, and was arguably mentally disturbed. Id. at 781–83, 743 P.2d at 236–37. In dissent, Chief Justice Pearson, after comparing Rupe’s case to others involving first-degree robbery and multiple victims, concluded:

Rupe’s case is the sole case in which the victims did not suffer prior to death and the defendant acted under the influence of an extreme mental disturbance. Just as the murders in this case were wanton and freakish in light of Rupe’s entire personal history, so the imposition of the death penalty is wanton and freakish in light of the other defendants.

Id. at 788, 743 P.2d at 239 (Pearson, C.J., dissenting) (emphasis in original).
326. Id. at 770, 743 P.2d at 230.
Conduct of Comparative Proportionality Review

sentence imposed on Brian Lord in *State v. Lord*. Here, the court first reiterated its belief that the purpose of comparative proportionality review is to alleviate “the types of major systemic problems identified in *Furman*: random arbitrariness and imposition of the death sentence based on race.” It then rejected any approach that seeks to achieve these ends by simply comparing “numbers of victims or other aggravating factors” that may “superficially make two cases appear similar” (as the court itself had done in previous years), and it did so on the ground that capital crimes are “unique and cannot be matched up like so many points on a graph.” Recognizing, though, that the very project of proportionality review requires some articulated basis for

327. 117 Wash. 2d 829, 916, 822 P.2d 177, 226 (1991). One year after *Rupe*, the court elaborated its criteria of similarity in *State v. Rice*, 110 Wash. 2d 577, 757 P.2d 889 (1988). In *Rice*, the court announced that it would consider the “heinous nature of [the] crimes, the number and severity of the aggravating factors, and the number of . . . victims.” *Id.* at 628, 757 P.2d at 917. However, in applying this expanded set of criteria, it concluded that the most relevant cases for purposes of comparison were *Jeffries, Campbell*, and *Rupe*. *Id.* at 625–27, 757 P.2d at 915–16. Citing these highly questionable precedents to justify its decision in *Rice*, the court thereby piled problematic precedent upon problematic precedent, rendering it ever more difficult for a future defendant to mount a successful proportionality challenge. Note, incidentally, that the court used *Rice*, prior to reviewing his case, to justify *Rupe*, and then used *Rupe* to justify its decision in *Rice*. *Rice*, 110 Wash. 2d at 625–26, 757 P.2d at 915–16; *Rupe*, 108 Wash. 2d at 768, 743 P.2d at 229. From this sort of circular reasoning, citing two cases to render each other proportionate, it would appear that no exit is possible. Once again, it is difficult to escape the conclusion that in this instance, as in others, the court engaged in selective identification of cases in order to sustain a particular outcome. This conclusion seems all the more plausible given the court’s conclusion on three points. First, there was credible mitigating evidence introduced during the sentencing phase of Rice’s trial (specifically, a lack of criminal history and a history of mental illness). *See Rice*, 110 Wash. 2d at 627, 757 P.2d at 916. Second, the court’s examination of the trial judge reports turned up four other cases in which there was “credible evidence of a mental disorder or diminished mental capacity,” but in which the death penalty was not imposed. *Id.* Third, the relevance of these other four cases was dismissed on various grounds, including the relative youth of two of the other defendants, the severity of their respective mental problems, and the fact that, in one, only a single juror had voted against the death penalty. *Id.* at 627–28, 757 P.2d at 916–17. While the first and second points do arguably distinguish these cases from *Rice* (whereas the third is irrelevant to comparative proportionality review), they are nonetheless vexing in light of the court’s earlier refusal to distinguish between, for example, *Campbell* and *Rupe*. *Rupe*, 108 Wash. 2d at 770, 743 P.2d at 230. In the former, the trial court did not find any mitigating evidence credible, State v. *Campbell*, 103 Wash. 2d 1, 30, 691 P.2d 929, 945–46 (1984). But, in the latter, the court made extensive findings of such evidence. *Rupe*, 108 Wash. 2d at 780–83, 743 P.2d at 235–37. If that difference is deemed of insufficient relevance in determining comparative proportionality, then it is hard to know why the relatively minor differences identified in *Rice* should be deemed sufficient to distinguish his case from others in which the death penalty was not imposed.


329. *Id.* (quoting *In re Jeffries*, 114 Wash. 2d 485, 490, 789 P.2d 731, 736 (1990)).

330. *Id.*
comparison, the court proceeded to invoke the work of the analytic philosopher Ludwig Wittgenstein in defense of what it called a “family resemblance approach”: 331

Although the cases where death was imposed do not necessarily have one characteristic or set of attributes in common, we nonetheless recognize that they are somehow related. This relation cannot easily be described; it consists of a complicated network of overlapping similarities—much like members of the same family, who can be recognized as relatives, even though they do not all share any one set of features. Thus, we examine prior cases for those which belong together because they resemble each other. 332

It is not clear that a defendant sentenced to death should find it reassuring when a State Supreme Court indicates that its proportionality review will be predicated on what it calls the “impressionistic” identification of unspecified resemblances that render various cases somehow related. 333 Perhaps sensing the inadequacy of this appeal, the court then proceeded to indicate the resemblances it now deemed most germane: the nature of the crime committed, the aggravating circumstances found by the jury, the defendant’s personal history, and, finally, the defendant’s criminal record. 334 Aside from rendering the review process ostensibly more sophisticated in a theoretical sense, it does not appear that the court’s appropriation of one of the central metaphors of an Austrian philosopher effected any substantive changes in its method of folding these criteria of similarity into judgments of proportionality. As in the past, the court justified its affirmation of Lord’s death sentence by citing the results of its previous reviews, with the aim of showing that his crime, character, and record were more heinous than those of his predecessors. Yet, in making this argument, the court cited only those cases that served its present purpose (specifically, those characterized by deaths that involved neither torture nor protracted suffering—e.g., Rupe, Harris, and Jeffries), while those that might render its conclusion problematic by virtue of their greater brutality (e.g., Campbell’s revenge murder of two women and a child) made no

331. Id. at 911, 822 P.2d at 223; see Ludwig Wittgenstein, Philosophical Investigations §§ 65–67 (1958).
332. Lord, 117 Wash. 2d at 911, 822 P.2d at 223.
334. Id.
Conduct of Comparative Proportionality Review

appearance in the court’s opinion. 335
The court offered its most ambitious pre-1995 effort to make good on its “family resemblance” approach when, in 1993, it affirmed the death sentence imposed on Gary Benn.336 To begin, the court isolated thirty trial judge reports regarding defendants convicted of murdering two or more persons, as was Benn,.337 Following a bare-bones recitation of the content of these reports, as well as a still more abbreviated review of the arguments advanced by the prosecution and defense regarding the cases each believed were or were not comparable, the court once again acknowledged the “difficulties inherent to the identification of ‘similar cases.’” 338 That noted, the court then removed most cases from this pool without, in a majority of instances, explaining why it did so (although it did suggest that, more often than not, the factors that appear to exempt defendants from the death penalty include mental disturbances, guilty pleas, and/or youthfulness).339 What remained following this process of

335. In order to discount the significance of “similar” cases in which the death penalty was not sought, with reference to each of the four prongs of its new test, the court simply cited one or two trial judge reports without any elaboration. Lord, 117 Wash. 2d 829, 911–14, 822 P.2d 177, 224–25. For example, with regard to its analysis of the nature of the crime, the court cited two trial judge reports which, like Lord, involved only a single victim, but which, unlike Lord, involved defendants with no prior convictions for a violent crime. Id. at 911, 822 P.2d at 224. The court thereby ignored trial judge reports filed in other rape/murder cases involving at least three aggravating factors, as was the case with Lord. Id. at 911–12, 822 P.2d at 224. On this point, see Justice Utter’s dissent in State v. Lord. In Lord, Justice Utter stated:
The majority focuses on those cases in which the death penalty was actually imposed. It expends precious little ink in describing a few similar cases in which the death penalty was either not imposed or not sought by the prosecutor. It does not list all of the similar cases. When the majority does mention such cases, it does not describe the aggravating factors and mitigating factors in those cases. It does not describe the defendant’s prior convictions in those cases. It superficially distinguishes those cases from this one without meaningfully comparing all of the relevant factors in those cases to the ones in this case. The majority then concludes that Lord’s sentence is proportionate, because, in the majority’s view, the death penalty has been imposed in similar cases.
Id. at 939–40, 822 P.2d at 239 (Utter, J., dissenting). Justice Utter then proceeded to identify five cases, all involving rape/murder and at least three aggravating circumstances, in which the death penalty was not imposed by the jury or in which the state never sought the death penalty. Id. at 943, 822 P.2d at 241 (Utter, J., dissenting). In addition, he identified four more cases, which, although not involving the same number and type of aggravating circumstances, were nonetheless similar in terms of the nature of the crime, the existence of a prior record, and the absence of credible evidence of mitigating circumstances. Id. at 944–45, 822 P.2d at 241–42 (Utter, J., dissenting). In each of the four, either the state did not seek the death penalty or the jury did not impose it. Id.
336. Benn, 120 Wash. 2d at 680, 845 P.2d at 317; see also Gilbert, supra note 217, at 607–12 (providing a detailed analysis of the court’s review in Benn).
338. Id. at 691, 845 P.2d at 323.
339. Id. at 692, 845 P.2d at 323. In his analysis of its review of Benn’s sentence, Gilbert argued
exclusion was a pool of seven defendants, four of whom had been sentenced to life imprisonment without possibility of parole and three of whom had been sentenced to death. The three sentenced to death, however, included one defendant whose case had never been reviewed by the court (Clark Hazen, who committed suicide while incarcerated), as well as two others (Patrick Jeffries and Mitchell Rupe) whose reviews were suspect for the reasons indicated above. Relying on this unreliable group, the court affirmed Benn’s sentence on the ground that this group “does not contain an arbitrary frequency of life without parole sentences over death sentences.” As before, the court justified its conclusion by citing its deference to the jury:

This court may systematically seek to undo and thus eradicate arbitrariness in sentencing. It will have limited success, however, in systematizing the unpredictable impulse toward mercy among juries which must decide cases that are ultimately as unique as each defendant. We have not sought to substitute this court’s judgment for that of the jury.

In addition to granting the impossibility of conducting comparative proportionality review in a way that satisfies its designated purposes, this conclusion appears to suggest the court’s willingness to affirm the jury’s determination in any given case, so long as it can find a small handful of “similar” cases that resulted in death sentences, no matter how disproportionate they may be when compared to the overall universe of cases.

that the court was inconsistent in excluding from the pool of similar cases those that exhibited one or more of these three factors. Gilbert, supra note 217, at 615. For example, the court excluded young defendants who did not receive the death penalty from its pool of similar cases. Id. Yet, without explanation, it then included in its pool of similar cases an eighteen-year old (Clark Hazen) who was sentenced to death. Id. Gilbert concluded: “The only plausible reason that the Hazen case was included by the majority was to make Benn’s sentence appear more proportionate.” Id.

340. Id. at 692, 845 P.2d at 323.
341. Id. at 684 n.9, 845 P.2d at 319 n.9 (stating that Clark Hazen committed suicide in prison before his conviction or sentence could be reviewed).
342. Benn, 120 Wash. 2d at 692, 845 P.2d at 323.
343. Id. at 692–93, 845 P.2d at 324.
344. Registering this concern in his dissent, Justice Utter reviewed fifteen cases not deemed similar by the court (including two the court did not list in its pool of thirty), none of which resulted in a death sentence. Id. at 700, 845 P.2d at 328 (Utter, J., dissenting). Adding these cases to those the majority found similar, Justice Utter determined that the death penalty is imposed on members of this pool of twenty-seven at a rate of approximately 11%. Id. at 706, 845 P.2d at 330 (Utter, J., dissenting). That, he concluded, does not begin to approach the requirement that death sentences be affirmed only if this penalty is imposed “generally” in similar cases. Id. at 709, 845 P.2d at 332.
Conduct of Comparative Proportionality Review

In sum, during its first decade of comparative proportionality review the Washington State Supreme Court did in time correctly conclude that, at least in principle, RCW section 10.95.130 requires it to include in its universe all cases in which defendants were convicted of aggravated first-degree murder, regardless of whether the death penalty was or was not sought or imposed. Turning from the universe of cases to the pool of similar cases to be abstracted from that universe, over the course of its first decade the court moved from its initial position, in which it essentially abandoned the project of comparative proportionality review in favor of inherent review, to its eventual articulation of Lord’s four criteria of similarity. Yet, as the cases reviewed above indicate, the court’s actual selection of cases deemed similar was unsystematic at best and, at worst, chosen with an eye to a pre-determined outcome. Finally, in terms of the third element of the logic of comparative proportionality review, the actual disposition of these cases, to all appearances the court selectively cited those elements of previously affirmed cases that would warrant an affirmation of proportionality, while ignoring those that might complicate or contradict that judgment.

2. Comparative Proportionality Review After 1995

In 1995, at the close of the Washington State Supreme Court’s first decade of comparative proportionality review, Justice Utter concluded that the court’s conduct of such review “increases rather than decreases the likelihood the death penalty will be imposed in an arbitrary and standardless manner, in violation of the equal protection clause of the fourteenth amendment to the United States Constitution.” Justice Utter’s criticism echoed a federal district court opinion from the previous year, which held that the Washington State Supreme Court failed to “fulfill the essential function of ensuring the evenhanded, rational, and consistent imposition of death sentences under Washington law.” Although the supreme court would eventually disavow this

(Utter, J., dissenting). “I cannot,” he closed, “ease the requirements of comparative proportionality review to the point where it becomes an empty ritual.”

346. Harris v. Blodgett, 853 F. Supp. 1239 (W.D. Wash. 1994), aff’d, 64 F.3d 1432 (9th Cir. 1995). Among other issues, this habeas petition raised the question of whether the Washington State Supreme Court had conducted an adequate proportionality review in Harris’s case. Id. at 1286. Answering this question in the negative, the federal court identified five flaws in the court’s practice. First, although RCW 10.95.130 specifies where to find similar cases, it does not define what counts as a similar case, and the court has failed to clarify how this determination should be
criticism, in a trio of death sentences affirmed in 1995 it nonetheless took the district court’s ruling as an opportunity to assess and reconsider its prior approach to the conduct of such review. Together, these decisions established the basic template for the court’s conduct of comparative proportionality review since that date. However, they also display the ongoing defects of that conduct.

As the law requires, the court since 1995 has remained formally committed to a universe of cases that includes all convictions for aggravated first-degree murder, regardless of whether the death sentence was imposed or executed. Moreover, since 1995, in each of its reviews the court has applied a four-prong test in identifying a pool of similar cases and in determining whether the sentence, when compared to other members of that pool, is disproportionate. That test requires a consideration of the nature of the crime, the number of aggravating circumstances, the defendant’s criminal history, and, finally, the defendant’s past. The standardization of this test, however, has done little to remedy the arbitrary nature of its pre-1995 reviews, and so, today, the court remains unable to offer a meaningful account of why some are sentenced to die, while the vast majority are not.

In the first of these three 1995 cases, State v. Gentry, the court once again applied the four family resemblances (nature of the crime, made. Id. at 1288. Second, when conducting any given review, there is no procedure for the parties to be notified of which cases the court considers similar until they receive its ultimate determination. Id. at 1286. Third, when no similar case can be identified, as in Campbell, no alternative procedure is available to inform the conduct of comparative proportionality review. Id. at 1289. Fourth, the court has not indicated the standards it employs in assessing proportionality through reference to the cases ultimately selected for the purposes of comparison. Id. Finally, the court presents information from the trial judge report filed in Harris’s case, as in other cases, as findings of fact, although “it is anyone’s guess whether the Report of the Trial Judge under RCW 10.95.120 is intended to be findings of fact.” Id. at 1290.

347. The Washington State Supreme Court indicated that because the Ninth Circuit, in affirming Harris, did not reach the question of proportionality, it did not consider itself bound by the district court decision. In re Benn, 134 Wash. 2d 868, 928, 952 P.2d 116, 147 (1998) (“The existence of an analytically flawed federal district court decision is not a compelling reason to vacate this defendant’s death sentence or reconsider the proportionality review in his case.”).

348. Brett, 126 Wash. 2d at 207–09, 212, 892 P.2d at 66–67, 69 (“Without determining the merits of these challenges under this court’s current proportionality review, we take this opportunity to revisit the development of such review in Washington and evaluate the continued viability of our present approach.”); see State v. Pirtle, 127 Wash. 2d 628, 683, 904 P.2d 245, 275 (1995); State v. Gentry, 125 Wash. 2d 570, 654–58, 888 P.2d 1105, 1154–56 (1995).

349. WASH. REV. CODE § 10.95.130(b) (2000).


Conduct of Comparative Proportionality Review

aggravating circumstances, prior convictions, and personal history) articulated in Lord. However, its consideration of the single aggravating circumstance consisted of nothing more than citation of the relevant statutory category (the crime was committed to protect the identity of the defendant); and its consideration of the nature of the crime consisted of the claim that Gentry’s offense was at least as brutal as that of Lord, which, because it culminated in an affirmed death sentence, rendered Gentry’s case not disproportionate. No other cases that might have complicated this conclusion were cited or considered.

With respect to the third family resemblance, the defendant’s prior record, the court merely cited Gentry’s previous convictions, but offered no consideration of any other cases, thereby rendering it impossible to decipher the sense in which this constituted a specifically comparative review. Finally, with respect to the defendant’s personal history, the court offered a bare bones citation of six trial judge reports, including those for David Rice and Westley Allan Dodd, in which evidence of mitigating circumstances such as youth, child abuse, or mental illness had not been sufficient to exempt defendants from a sentence of death.

In the second of these three cases, State v. Brett, the court once again returned to its view that the purpose of the proportionality review provision of RCW section 10.95 is not to ensure statewide consistency in capital sentencing, but merely to identify cases that are grossly disproportionate. The justification it proffered for this view was new, however. Specifically, the court stated that the legislative provisions intended to channel jury discretion (e.g., the requirement that at least one aggravating factor be found applicable) “ensure proportionality and eliminate the ability of the jury, in all but the most aberrant case, to impose the death sentence in a wanton and freakish manner.”

352. Id. at 656, 888 P.2d at 1155.
353. See id. at 656–57, 888 P.2d at 1155–56.
354. See id. at 657, 888 P.2d at 1155–56.
355. Id. In his dissenting opinion in Gentry, Justice Utter indicated that an examination of the trial judge reports turned up at least twelve cases involving multiple murders (which Gentry’s did not), murders involving extreme and prolonged suffering on the part of the victim (which Gentry’s may or may not have, depending on how one reads the autopsy evidence), and a victim who was particularly vulnerable by virtue of age (as was the case with Gentry). Id. at 669–71, 888 P.2d at 1161–63 (Utter, J., dissenting). On the face of it, all of these murders appear more brutal than that committed by Gentry, and yet in none was the death penalty imposed. Id.
357. See id. at 212–13, 892 P.2d at 68–69.
358. Id. at 210–11, 892 P.2d at 68.
court thus came very close to rendering its own reviews superfluous. By this account, except in extraordinary instances, the legislatively prescribed procedures that structure jury deliberation are presumed to ensure against disproportionality.359

Having thus (mis)construed its task, the court then quite remarkably repudiated any attempt to extract a smaller pool of similar cases from the universe of cases available to it based on the number or kind of aggravating factors, the “family resemblances” articulated in Lord, or, for that matter, any other criterion. Instead, conflating the distinction between universe and pool, the court defined as “similar” all cases for which trial judge reports had been filed or cases reported in the Washington Reports or Washington Appellate Reports since January 1, 1965.360 In defense of this violation of the logic of comparative proportionality review, the court contended that this definition remedied one of the due process concerns expressed by the federal district court.361 Specifically, by conflating the distinction between universe and pool, the court rendered it impossible for a defendant to claim that he or she could not know the criteria employed in identifying the cases the court found comparable to his or hers (or, it might be added, to contest those criteria as well as their application in isolating some smaller pool from the larger universe of cases).362 Finally, absent any elaboration of what it had indeed found in this larger universe, and absent any discussion of specific cases other than that of Brett, the court declared that “after carefully reviewing the totality of similar cases,” it deemed Brett’s case “not disproportionate”: “There is no unique or distinguishing characteristic of the Defendant or of this crime which makes imposition of the death penalty wanton and freakish.”363 In sum, after discarding all
Conduct of Comparative Proportionality Review

previous attempts to render this process meaningful, the court abandoned any effort to articulate a method to inform its deliberations and, instead, opted to substitute an uncontestable declaration of its unexplained affirmation.

In affirming the death sentence imposed on Blake Pirtle, the final member of the trio of cases reviewed in 1995, the court consolidated the basic template for all reviews conducted since that date. In its opinion, the court essentially split the difference between Lord and Brett, and it did so in a way that, arguably, combined the most troublesome aspects of each. Citing Brett, the court in Pirtle restated its assumption that juries will not render aberrant judgments; and, again citing Brett, it reaffirmed its commitment to consider as similar all cases specified by RCW section 10.95.130(2)(b). Although rejecting the “family resemblance” test articulated in Lord on the grounds that it had proved “somewhat unwieldy as more and more cases were reported,” the court nonetheless adopted the very same “resemblances” it had identified as most salient in that case. Specifically, and without explaining why these criteria and only these merit consideration, the court stated that in conducting comparative proportionality analyses, it would assess the nature of the crime, the number of aggravating circumstances, the defendant’s criminal history, and, finally, the defendant’s past.

vacillation on this question. Because virtually every review conducted prior to and including Brett was decided on a novel basis, the authority of all these judgments is compromised. This has potentially lethal consequences for those sentenced to death. See id. at 214–16, 828 P.2d at 70–71 (Durham, J., concurring) (“In virtually every recent death penalty case decided by this court, a different definition of proportionality has been promulgated. The resulting confusion is unacceptable” because, arguably, any given determination of “proportionality itself depends on the consistency of a proportionality analysis.”).

364. Justice Utter stated:
Even if ‘aberrant’ or ‘wanton and freakish’ were the standard, and it is not, it is impossible to conclude a given sentence is not ‘aberrant’ or ‘wanton and freakish’ without engaging in some process of reflection, whatever that may be. The designated majority requires that one simply take its word on so important a question as whether a defendant properly may be executed, without revealing what that process is. It thus forecloses any possibility of review or even discussion of its conclusion.

See id. at 230, 828 P.2d at 77 (Utter, J., dissenting).
366. See id. at 683–89, 904 P.2d at 275–77.
367. Id. at 686, 904 P.2d at 276.
368. Id.
369. Id. at 687, 904 P.2d at 276.
370. Id. at 686, 904 P.2d at 276.
Applying these criteria, the Pirtle court concluded that “there is no factor or combination of factors which marks this as an unusual, let alone wanton or freakish, death penalty case.” 371 Although the court cited no specific case in defense of this conclusion, it did note, first, that only 20% of the cases on file involved, as did Pirtle, three or more aggravating factors; second, that approximately one-third involved, as did Pirtle, more than a single victim; and, finally, that only 3% of the cases involved, again as did Pirtle, defendants with more than ten prior convictions. 372 It is difficult to know how much weight to ascribe to these quantitative determinations because the court also reiterated its oft-made claim that the determination of proportionality is not “a statistical task” that “can be reduced to a number,” and that “numbers can” only “point to areas of concern.” 373 Yet when it considered those matters that

371. Id. at 688, 904 P.2d at 277.
372. Id. at 687–88, 904 P.2d at 277.
373. Id. at 687, 904 P.2d at 277. Were the court to engage in a serious effort at frequency analysis, it would almost certainly discover several areas of significant concern. First, for example, all other things being equal, one would assume that the greater the number of victims in any given case, the more likely it is that the death penalty will be sought. Indeed, it is true that of the total number of convictions for aggravated murder, the death penalty has been sought in 27% of the cases involving a single victim; 34% when there were two; and 47% when there were three. However, the death penalty was sought in only 20% of the cases involving four victims, never in those involving five victims or, for that matter, in the lone case involving ten. Granted, the death penalty was sought for two co-defendants convicted of jointly murdering thirteen persons (although it was imposed on only one of the two, Kwan Mak, and his sentence was subsequently set aside). Mak v. Blodgett, 970 F.2d 614 (9th Cir. 1992). However, in 2000, Robert Yates evaded a likely death sentence in Spokane County after confessing to the murder of thirteen victims, only to be sentenced to death two years later in Pierce County for the killing of two. Christine Clarridge, Death Sentence for Yates Elicits Tears, but None for Joy, SEATTLE TIMES, Oct. 4, 2002, at A1. Finally, the sole person convicted of murdering a larger number, Gary Ridgway, was able to escape an almost certain death sentence by pleading guilty to killing forty-eight women. Kershaw, supra note 4, at A1. If nothing else, these calculations suggest that body count cannot be taken as a reliable predictor of whether or not a defendant will be sentenced to death, which implies in turn that those responsible for a far smaller tally cannot be confident of escaping the death chamber.

Second, again all other things being equal, if the court takes the number of aggravating factors as an indicator of the heinousness of any given crime, as the court clearly does, one would expect that the percentage of death sentences sought and imposed would increase as the number of aggravating circumstances rises. However, that does not consistently hold true. Setting aside the cases in which the number of aggravating circumstances is not indicated in the trial judge reports, 47% of the remaining convictions for aggravated first-degree murder involved a single aggravating factor. In 25% of these cases the death penalty was sought, and in 5% it was imposed. In cases involving two aggravating factors (29% of the total), the rate at which the death penalty was sought remained the same (25%), but the rate at which it was imposed rose from 5% to 28%. In cases involving three aggravating factors (15% of the total), the rate at which the death penalty was sought jumped from 25% to 47%, and the rate at which it was imposed increased from 28% to 37%. To this extent, these numbers correspond, at least in rough terms, to what one might reasonably anticipate. However, when one considers the cases involving four aggravating factors (6% of the total), the death penalty
Conduct of Comparative Proportionality Review

can be quantified, with the exception of Pirtle’s crime, which it described in a brief paragraph, it offered nothing other than the bare figures noted here. Because one can only guess which cases provided the basis for these percentages, these figures are impossible to challenge.

Turning to matters that are not susceptible to quantitative analysis, by refusing to cite specific cases the court in Pirtle avoided the comparability challenges that might otherwise follow. But this omission also renders it impossible to fathom the basis for its affirmation of this sentence. This is most readily apparent when the court considered Pirtle’s character, which it described through reference to the mitigating evidence introduced during the sentencing phase of the trial. Noting that Pirtle proffered evidence of drug abuse, alcohol addiction, and an abusive family history, the court contended that “an examination of the aggravated murder reports shows that some mitigators—for instance, mental illness or extreme emotional distress—may lead prosecutors and juries not to seek or impose the death penalty. Neither addiction nor a history of abuse as a child appears to have such an effect.” It is difficult to know how the court can sustain this claim given the extreme poverty of the trial judge reports, especially when it comes to their accounts of mitigating circumstances. Moreover, assuming that the information contained in these reports is correct, its claim in this regard is simply wrong. Addiction and childhood abuse are in fact often offered as credible mitigating evidence in cases that have not generated death sentences, and so it is impossible to know what led the court to the

was sought in a whopping 73% of the cases, but the rate of imposition dropped from 37% to 27%. Finally, when one considers the very small number of cases involving five or more aggravating factors, the rate at which the death penalty was sought decreased significantly (from 73% to 43%), but its rate of imposition increased dramatically (from 27% to 67%).

Of course, one might contend that the court’s overall rationale in the cases it has reviewed for proportionality becomes explicable not when one considers any of the four prongs of the Pirtle test individually, but only when one considers their cumulative combination in any given case. In order for that claim to prove credible, however, the court would have to abstract from the four prongs of its current test some generalized measure of defendant culpability, and then compare each defendant’s degree of culpability with that of others, again using the same abstracted standard in order to avoid the problems that crop up when the court seeks to compare cases on the basis of factual similarity. That, however, the court has not done.

374. Pirtle, 127 Wash. 2d at 688, 904 P.2d at 277.
375. Id.
376. Cases considering these factors include: Report of Trial Judge No. 8 (Charles Bingham); No. 20 (William Martin); No. 44 (Dennis Williams); No. 45 (James Dykgraaf); No. 50 (Sean Stevenson); No. 52 (Christopher Blystone); No. 53 (James Thompson); No. 58 (Gene Kane); No. 64 (Jonathan Woods); No. 65 (Jeffrey Lane); No. 68 (Darrin Hutchinson); No. 77 (Charles Tate); No. 80 (Gabriel Garcia); No. 93 (Timothy Caffrey); No. 95 (Kenneth Schrader); No. 182 (Joey Ellis); No. 186 (Gerald Davis). These reports expressly cite alcohol and/or childhood abuse as credible
conclusion that Pirtle’s “mitigation evidence was, at best, average.” Still more troubling, once the court recorded this conclusion, it then began to cite Pirtle as precedential authority in affirming the same conclusion under equally suspect circumstances.

Since Pirtle, the court has applied this four-prong test to each of the death penalty cases it has reviewed on direct appeal. The court has vacillated, however, with respect to its refusal, in Brett as well as Pirtle, to cite specific cases to justify its conclusions. When it has elected to do so, the court typically cites only those cases that did in fact culminate in death sentences and that were subsequently reviewed and affirmed.

It thereby violates what it expressly affirmed in Rupe: the Washington statute requires consideration of all convictions for the crime of aggravated first-degree murder, including those in which the death penalty was not sought. That in turn cannot help but bias its review toward affirmation. For example, in State v. Brown, the court first cited the affirmed death sentences imposed on Westley Allen Dodd, Mitchell Rupe, Patrick Jeffries, Gary Benn, Benjamin Harris, and Jonathan Gentry in order to sustain its claim that Brown’s crime was at least as heinous as those committed by these defendants. Second, it cited the cases of Gentry, Benn, Harris, and Jeffries in order to justify its claim that the number of aggravating factors found applicable in

mitigating evidence in the cases of defendants not sentenced to death. It is true that the mere citation of such mitigating evidence does not allow one to know for certain that the jury was influenced by that evidence when rendering its sentencing decision. However, the Washington State Supreme Court is in no position to know this either.

377. Pirtle, 127 Wash. 2d at 688, 904 P.2d at 277.
378. See, e.g., State v. Elmore, 139 Wash. 2d 250, 310, 985 P.2d 289, 323 (1999) (holding that Clark Elmore’s personal history did not excuse his crime despite his history of abuse as a child, the court relied on State v. Brown, 132 Wash. 2d 529, 940 P.2d 546 (1997), which, in turn, was based on Pirtle’s erroneous conclusion).
381. Cases in which comparative proportionality review is limited to affirmed death sentences include: Woods, 143 Wash. 2d at 616, 23 P.3d at 1076; Davis, 141 Wash. 2d at 880, 10 P.3d at 1023–24; Elmore, 139 Wash. 2d at 308, 985 P.2d at 322; Sagastegui, 135 Wash. 2d at 92, 954 P.2d at 1324; Brown, 132 Wash. 2d at 555–56, 940 P.2d at 562.
383. Id. at 557 n.56, 940 P.2d at 562 n.56.
Conduct of Comparative Proportionality Review

Brown’s case was consistent with (or in some cases less than) the number in these other cases. Third, it cited Rupe, Harris, and Benn in order to show that Brown’s prior record was more extensive than these other defendants. Finally, it cited Dodd, Brian Lord, and David Rice in order to show that the sort of mitigating evidence adduced in Brown’s case (childhood abuse and various mental as well as personality disorders) has not precluded the court from affirming death sentences in the past. Consideration of each of the four prongs of the Pirtle test consumes a single brief paragraph, and none involves any comparative analysis of the companion cases it cited or any treatment of aggravated murder convictions that did not result in death sentences.

On other occasions, the court has not cited specific cases in its reviews, which renders it difficult at best to know on what basis the court rests its conclusions. For example, in State v. Stenson, the court stated that “this case involved a greater degree of premeditation than in many other cases of first-degree aggravated murder.” However, because no other cases are cited, it is impossible to grasp the rationale for this inherently comparative judgment. The court’s failure becomes still more glaring when, in considering one or more prongs of the Pirtle test, it offers no comparative reference whatsoever, instead limiting itself to a mere restatement of the facts immediately relevant to the prong in question. Again using Stenson as an example, in considering the relevant mitigating circumstances, the court’s “analysis” consisted of the following statement: “The Defendant did have some criminal history, including felony drug convictions, although none of the prior crimes were crimes of violence.” The qualification registered at the end of this sentence is precisely the sort that should in principle invite rigorous examination of the defendant’s character and biography in order to determine whether the sentence imposed in his case was

384. Id. at 558 n.60, 940 P.2d at 563 n.60.
385. Id. at 558 n.62, 940 P.2d at 563 n.62.
386. Id. at 559 n.66, 940 P.2d at 563 n.66.
387. Id. at 556–59, 940 P.2d at 562–63.
389. Id. at 759, 940 P.2d at 1285.
390. Id. at 759–60, 940 P.2d at 1285.
391. Id. at 760, 949 P.2d at 1285. By the same token, in considering the applicable aggravating circumstances in this case, the court’s review consisted of the following: “The aggravating circumstances are that there was more than one murder victim and that Frank’s murder was committed to conceal the identity of Denise’s murderer.” Id. at 759–60, 949 P.2d at 1285.
disproportionate. The court, though, offered no such analysis, thereby
making a mockery of the purpose of comparative proportionality review
as well as its own four-prong test.

The Washington State Supreme Court is currently in a position to
affirm any death sentence that comes before it by picking and choosing
among the death penalty cases it has previously reviewed and affirmed.
While the charade of the four-prong test articulated in *Pirtle* is formally
perpetuated, perhaps a more candid statement of the court’s true test was
advanced in 1999 when it affirmed the death sentence imposed on Clark
Elmore: “If the facts of Elmore’s case are similar to some of the facts
taken from cases in which the death penalty was upheld, the
proportionality review is satisfied.” To all appearances, the court now
finds application of this test virtually effortless. Should, for example, a
defendant sentenced to death have little or no prior criminal record,
Mitchell Rupe’s case can be pressed into service in order to show that
the court has in the past upheld death sentences imposed on defendants
who had no prior record (which renders it impossible to ever generate
the conclusion that a sentence is disproportionate on this basis).

Should a defendant have a history of diagnosed mental disorders, the
cases of David Rice and Westley Allan Dodd can be trotted out in order
to demonstrate that the court in the past has upheld death sentences
imposed on defendants who were at least as disturbed as the present
defendant. Should a defendant be sentenced to death on the basis of
two aggravating factors, the court can put to work the cases of Jonathan
Gentry, Gary Benn, and Benjamin Harris to show that the court has
upheld sentences when only a single aggravator was found applicable, or,
if only a single aggravator is found applicable, the court can cite
*Gentry*, *Benn*, and *Harris* once again in order to show that a case under
review is not disproportionate.

the court affirmed the death sentence imposed on Rupe even though he, like Sagastegui, had little or
no criminal record).
(stating that the court has upheld death sentences imposed on defendants, including Dodd and Rice,
who suffered from personality disorders similar to, or more severe than, that suffered by Brown).
395. See, e.g., Elmore, 139 Wash. 2d at 309, 985 P.2d at 323 (stating that the court has upheld
death sentences imposed on defendants, including Gentry, Benn, and Harris, when only a single
aggravator was found applicable).
396. See, e.g., State v. Elledge, 144 Wash. 2d 62, 81, 26 P.3d 271, 282 (2001) (stating that the
court has upheld death sentences imposed on defendants, including Gentry, Benn, and Harris, when
Conduct of Comparative Proportionality Review

True, the precise cast of characters alters slightly from one review to another, if only in the sense that each time another death sentence is affirmed an additional name becomes available to cite in authorizing subsequent affirmations. That said, the script now changes only in order to accommodate each new defendant, the plot unfolds with mechanical uniformity, and the dreary dénouement never varies. Even when analysis is confined to the small number of death penalty cases reviewed by the Washington State Supreme Court since 1981, it is impossible to construct a reasoned account that explains why all were deemed sufficiently proportionate to warrant affirmation of their respective death sentences. A less flawed comparative proportionality analysis would require a far more sophisticated method than the court currently employs as well as a far more detailed analysis of death sentence cases than it currently provides. Moreover, even if the court were to commit itself to such a method, it is not clear how it could fulfill this pledge given the irremediable deficiencies of the trial judge reports on which it would have to predicate such an inquiry.

Taking note of the atrophy of the court’s comparative proportionality review into a pro forma ritual, shortly before he retired from the bench, Justice Utter offered an apt characterization of the history I have traced here.\textsuperscript{397} What the court in \textit{Brett} characterized as ‘an increasingly broad approach’ to defining ‘similar cases’ is more aptly described as the gradual degeneration of judicial review in capital cases, a process which reaches its low point with the introduction into our proportionality analysis of a new, and curiously elusive, concept: all murders falling within the purview of RCW 10.95 are, ipso facto, proportionate—except when they are not.\textsuperscript{398}

Except, Justice Utter might have added, that they are never in fact not.\textsuperscript{399}

---

398. \textit{Id.}
399. The sophistry apparent in the court’s conduct of comparative proportionality reviews has apparently proven too much for at least one of its current members to bear. Thus, in \textit{Elledge}, Justice Sanders condemned the court for initiating its review in any given case on the basis of the premise that “the defendant is ‘qualified’ for the death penalty so long as it is not ‘wantonly and freakishly’ imposed, notwithstanding how many others may have engaged in similar conduct who were not executed.” 144 Wash. 2d at 88, 26 P.3d at 285 (Sanders, J., dissenting) (emphasis in original). The statutory requirement of comparative proportionality review, Justice Sanders concluded, “has degenerated through numerous iterations into the current ‘wanton and freakish’ standard, finally becoming little more than lip service to the important protection proportionality review was
V. THE FAILURE OF COMPARATIVE PROPORTIONALITY REVIEW AND THE FUTURE OF THE DEATH PENALTY

In *Gregg v. Georgia*, Justices Stewart, Powell, and Stevens expressed considerable confidence in the capacity of the safeguards adopted by Georgia and affirmed by the U.S. Supreme Court in 1976 to prevent arbitrary and capricious imposition of the death penalty:

No longer can a jury wantonly and freakishly impose the death sentence; it is always circumscribed by the legislative guidelines. In addition, the review function of the Supreme Court of Georgia affords additional assurance that the concerns that prompted our decision in *Furman* are not present to any significant degree in the Georgia procedure applied here.  

Echoing this claim in *State v. Rupe*, the Washington State Supreme Court upheld Washington’s death penalty law on the basis of its confidence in these same safeguards: “Defendant’s arguments were addressed by the Supreme Court in *Gregg*, and we find that analysis equally applicable here.” However, as was the case in *Gregg*, the *Rupe* court upheld a statute that had just been adopted and rarely applied. Nearly three decades later, we are in a position to assess whether state courts are indeed providing the protection promised in *Gregg*. Reviewing the studies that have examined this question, James Acker and Charles Lanier concluded: “A wealth of empirical research now exists on the operation of capital punishment statutes, and much of that evidence suggests that the premises underpinning the *Gregg* decision are fallacious. Arbitrariness and discrimination continue to plague the administration of death penalty legislation.”

Evidence from the State of Washington with respect to the conduct of comparative proportionality review furnishes no reason to believe otherwise. If Acker and Lanier are correct, then there are good grounds for concluding that the death penalty, as currently applied in the State of Washington, is unconstitutional. First, because the concerns that led the

---

401. *Id.* at 206–07 (plurality opinion).
Conduct of Comparative Proportionality Review

U.S. Supreme Court to invalidate all death penalty statutes in effect in 1972 remain, equally if not more, valid today, Washington’s death penalty statute is unconstitutional under Furman v. Georgia. Neither Gregg nor Rupe overruled Furman, and Furman stands today for the proposition that a capital punishment statute, in order to conform to the imperatives of the Eighth Amendment, must achieve two objectives: it must limit imposition of the death penalty to the small group of defendants for which it is appropriate, and ensure that the members of this small group are selected rationally and consistently. Leaving aside the disproportionate application of the death penalty to those who murder white victims, as well as the failure of the death penalty’s infrequent imposition to deter, the lack of any meaningful way of distinguishing those who receive the death penalty from those who do not has not been remedied by the procedural reforms adopted by Washington in 1981. The administration of the death penalty in Washington does not ensure that the death sentence is restricted to the

404. 408 U.S. 238 (1972) (per curiam).
405. See id. at 294 (Brennan, J., concurring).
406. See generally LARRANAGA, supra note 293 (analyzing racial and geographical disparities in the administration of capital punishment in Washington). Larranaga found that in the majority of counties in the state (57%) death notices have never been filed, and that in 74% of those counties no death sentence has ever been imposed. Id. at 17. With respect to racial disparities, Larranaga concluded that since 1981, death has been imposed at a higher rate against African-Americans as compared to Caucasians. Additionally, death is sought and imposed at a significantly lower rate when the victim is African-American. And finally African-American defendants charged with killing a Caucasian victim have a significantly higher percentage of death notices filed and death sentences imposed.

Id. at 26. Moreover, at the time Larranaga completed his study, in a state with an African-American population of 3.2%, see United States Dep’t of Commerce, Bureau of the Census, Washington Quick Facts, at http://quickfacts.census.gov/qfd/states/53000.html (last visited June 30, 2004), five of the ten on death row were African-Americans, all but one of whom were convicted of killing white victims. These findings are especially troubling in light of the Washington State Supreme Court’s denial, in 1995, that race plays any role in determining who is and is not sentenced to death. See State v. Gentry, 125 Wash. 2d 570, 655, 888 P.2d 1105, 1154 (1995) (“In this case, there is no evidence that race was a motivating factor for the jury, and contrary to the Defendant’s suggestion, a review of the first degree aggravated murder cases in Washington does not reveal a pattern of imposition of the death penalty based upon the race of the Defendant or the victim.”); see generally David C. Baldus et al., Racial Discrimination and the Death Penalty in the Post-Furman Era: An Empirical and Legal Overview with Recent Findings From Philadelphia, 83 CORNELL L. REV. 1638 (1998) (providing a recent study dealing with racial disparities and the death penalty on a national level).

407. See Furman, 408 U.S. at 312 (White, J., concurring) (stating that “a major goal of the criminal law—to deter others by punishing the convicted criminal—would not be substantially served where the penalty is so seldom invoked that it ceases to be the credible threat essential to influence the conduct of others”).
most heinous criminals guilty of the most brutal murders, and, correlative, it fails to ensure that those who are convicted of murders as brutal as those committed by those who are in fact sentenced to death receive that same punishment. Because the Washington capital punishment statute operates in an arbitrary and capricious manner, it violates the Eighth Amendment and so cannot stand.

Second, even if an appeal to the Eighth Amendment proves unpersuasive, a meaningful comparative proportionality review is nonetheless required by article 1, section 14 of the Washington State Constitution, which prohibits the infliction of cruel punishment. The U.S. Supreme Court’s interpretation of the Eighth Amendment does not control the Washington State Supreme Court’s interpretation of article 1, section 14; and, in \textit{State v. Fain}, the court reaffirmed its holding that Washington’s prohibition of cruel punishment is broader than that provided by the Eighth Amendment. Moreover, in \textit{State v. Roberts}, the court reiterated its conviction that the imposition of a capital sentence is cruel if it is imposed without an individualized determination that the punishment is appropriate. That imperative is all the more pressing because, via adoption of RCW section 10.95, and in particular its requirement of proportionality review by the State Supreme Court, the legislature has given defendants a legitimate expectation that capital sentences will be imposed in a way that is fair. Specifically, it entitles them to the expectation that similarly situated defendants will be afforded the same punishment and that those who are not so situated will be punished differently.

Comparative proportionality review is the only mechanism, which, at least in principle, enables system-wide evaluation of jury decision-making in order to ensure that the death penalty is not applied in an

---

408. \textsc{Wash. Const.} art. I, § 14.
409. 94 Wash. 2d 387, 617 P.2d 720 (1980).
410. \textit{Id.} at 392, 617 P.2d at 723.
411. 142 Wash. 2d 471, 14 P.3d 713 (2000).
412. \textit{Id.} at 502, 14 P.3d at 731.
413. In this context, it is worth noting that in 1997 the state legislature passed a bill, S.B. 5093, 55th Leg., Reg. Sess. (Wash. 1997), that would have eliminated the requirement that the State Supreme Court conduct comparative proportionality reviews in conjunction with its mandatory review of all death sentences. \textit{Id.} That bill, however, was vetoed by Governor Gary Locke, who stated: “I am a strong supporter of the death penalty. However, I am also a strong supporter of fairness. The proportionality review has not yet resulted in the reversal of any death sentences. Nonetheless, I believe that it is an important safeguard.” Veto Message on S.B. 5093 (April 24, 1997).
Conduct of Comparative Proportionality Review

arbitrary and capricious manner or on the basis of impermissible discrimination. To ensure that such application does not occur, the court must compare each aggravated murder conviction against all others, and in order to do that its review must be based on complete and accurate information. However, given the deficiencies of the trial judge reports, combined with those evident in the court’s own conduct of comparative proportionality review, it cannot be maintained that individual defendants are presently protected from arbitrary, discriminatory, or otherwise unfair death sentences.414 As such, the death penalty in Washington violates the state constitution’s prohibition of cruel punishments.

Third, in *Hicks v. Oklahoma*,415 the U.S. Supreme Court held that when a state enacts a criminal statute setting out a procedure for the imposition of a particular penalty, a defendant has a “substantial and legitimate expectation” that he or she will be deprived of liberty only if the state complies with the procedural requirements of that statute.416 Accordingly, a defendant convicted of aggravated first-degree murder and sentenced to death in Washington has a due process right to appellate proportionality review in conformity with RCW section 10.95. True, as noted earlier, the U.S. Supreme Court held in *Pulley v. Harris*417 that the federal constitution does not require such review.418 That, however, is irrelevant to this argument. Because the Washington statute establishes the procedural protection of proportionality review, as the federal district court reiterated in *Harris v. Blodgett*,419 the State Supreme Court is obligated under the Fourteenth Amendment’s Due Process Clause to ensure that this analysis complies with the state statute and that it is performed in way that is meaningful.420 This is all the more

Since the death penalty is the ultimate punishment, due process under this state's constitution requires stringent procedural safeguards so that a fundamentally fair proceeding is provided. Where the trial which results in imposition of the death penalty lacks fundamental fairness, the punishment violates article 1, section 14 of the state constitution.

416. *Id.* at 346.
418. *Id.* at 41–43.
so when the state seeks to deliberately extinguish human life, which, the
U.S. Supreme Court has ruled, demands a heightened degree of
procedural due process scrutiny, or what Margaret Radin has called
“super due process.”421 In other words, when the potential harm to a
defendant extends to the taking of his or her life, concerns about the
potential for arbitrary state action are greatest, and so safeguards against
such arbitrariness must be the most scrupulous. However, for the reasons
indicated in this Article’s discussion of the State Supreme Court’s actual
conduct of comparative proportionality review, it cannot be maintained
that such review has in fact been conducted in a meaningful way; and,
for the reasons indicated in its discussion of Washington’s trial judge
reports and their deficiencies, even if the Washington State Supreme
Court were to seek to remedy the deficiencies of its conduct, it could not
do so.422

Arguably, if one could somehow overcome the problems posed by the
defects of the trial judge reports, and were the Washington State
Supreme Court to render its conduct of comparative proportionality
review less hollow than it presently is, certain constitutional infirmities
might be alleviated. Other state supreme courts have attempted to do so.
Most notably, in 1988 New Jersey appointed a special master, David
Baldus, who developed a complex statistical methodology in order to
determine the frequency with which death sentences are or are not
imposed on different groups of defendants whose overall level of
culpability is comparable.423 (This methodology was subsequently
modified as a result of recommendations advanced by a second special
master in 1999).424 However, rather than explore these efforts to

424. For an account of the modifications adopted in response to the second master’s report, see In re Proportionality Review Project, 735 A.2d 528 (N.J. 1999) and In re Proportionality Project (II), 757 A.2d 168 (N.J. 2000).
Conduct of Comparative Proportionality Review

rationalize a system that remains infected by the same problems identified more than three decades ago, this Article’s conclusion suggests that no matter how salutary such reforms might be, they should not distract us from an appreciation of the way in which comparative proportionality review discloses and compounds the fundamental dilemma on which the law has foundered since Gregg in its effort to develop a coherent jurisprudence of capital punishment. If comparative proportionality review, no matter how refined, cannot successfully overcome this dilemma, if such review simply rearticulates that dilemma in a new guise, then it would appear that Justice Blackmun was correct when he concluded that no amount of tinkering can salvage the machinery of state-imposed death.425

The key premise of contemporary capital punishment jurisprudence is that the

penalty of death differs from all other forms of criminal punishment, not in degree but in kind. It is unique in its total irrevocability. It is unique in its rejection of rehabilitation of the convict as a basic purpose of criminal justice. And it is unique, finally, in its absolute renunciation of all that is embodied in our concept of humanity.426

On the basis of this premise, in *Woodson v. North Carolina*,427 the U.S. Supreme Court rejected a statute that provided for a mandatory death sentence for specific offenses on the ground that such a law denies to the sentencing authority the discretion that is crucial if that authority is to take into account the individual character of the defendant and the particular circumstances of his or her offense.428 Such a statute “treats all persons convicted of a designated offense not as uniquely individual human beings, but as members of a faceless, undifferentiated mass to be subjected to the blind infliction of the penalty of death.”429 One implication of *Woodson* was elaborated two years later when, in *Lockett v. Ohio*,430 the Court held that because “[t]he need for treating each defendant in a capital case with that degree of respect due the uniqueness of the individual human being is so fundamental that the State cannot constitutionally fail to accord it.”431

---

428. *Id.* at 304 (plurality opinion).
429. *Id.*
of the individual is far more important than in noncapital cases," the
sentencer cannot be precluded from "considering any aspect of the
defendant’s character and record or any circumstances of his offense as
an independently mitigating factor." 431

Yet it is precisely this sort of unrestricted discretion that led Justice
Douglas in Furman to conclude that the death penalty, as then
administered by the states, violated the Eighth Amendment ban on
punishments that are selectively applied. 432 Mandatory death penalty
statutes of the sort invalidated by Woodson were found to be arbitrary
because they require all persons convicted of a given capital offense to
be sentenced to death, regardless of relevant factors such as past criminal
record or the likelihood of future criminal conduct. 433 Death penalty
statutes that grant unfettered discretionary authority to sentencers were
also found to be arbitrary because their actual operation renders it
impossible to fathom why some are sentenced to die while others are
not, and because they leave open the possibility that defendants may be
sentenced to die on the basis of legally irrelevant factors such as race. 434
While the Gregg plurality sought to channel juries’ discretion by
denying them the authority to deem certain crimes worthy of death,
thereby reducing the likelihood that individuals convicted of similar
crimes will receive different sentences, 435 the Lockett Court endowed
juries with the authority to render their sentencing decisions on the basis
of any and all mitigating factors they consider relevant, thereby
increasing the likelihood that individuals convicted of similar crimes
will receive different sentences. 436

As numerous commentators have remarked, the bulk of post-Furman
capital jurisprudence can be understood as an attempt to negotiate these
competing imperatives, 437 each of which expresses a rival conception of
fairness, neither of which we can reject in good constitutional
conscience. One conception requires us to take into account the
individual character and circumstances of particular defendants in

431. Id. at 605, 607 (plurality opinion).
432. Furman, 408 U.S. at 242 (Douglas, J., concurring).
433. Woodson, 428 U.S. at 303–05 (Stewart, J., concurring).
434. Furman, 238 U.S. at 253–57 (Douglas, J., concurring); id. at 309–10 (Stewart, J.,
concurring); id. at 313 (White, J., concurring).
436. Lockett, 438 U.S. at 605 (plurality opinion).
437. The best account of the U.S. Supreme Court’s attempt to negotiate this dilemma remains that
provided by Steiker & Steiker. See Steiker & Steiker, supra note 41, at 417–18.
Conduct of Comparative Proportionality Review

determining what penalty is appropriate, while the other requires us to impose the same penalty on defendants convicted of the same crime. The U.S. Supreme Court has itself noted that these competing imperatives “can be in some tension”\(^{438}\) and so can require “somewhat contradictory tasks.”\(^{439}\) Indeed, it is precisely the effort to navigate between this jurisprudential Scylla and Charybdis that ultimately led Justice Blackmun, nearing retirement from the Court, to declare that these two requirements are “not only inversely related, but irreconcilable”\(^{440}\) and, on that basis, to conclude that the death penalty cannot be administered in a way that comports with the federal constitution.\(^{441}\) If the operation of capital punishment’s competing legal imperatives ensures that disparate sentencing for defendants who commit comparable crimes is inevitable, then arbitrariness in the imposition of the death penalty is inevitable as well.

The fundamental conceptual and practical problems involved in comparative proportionality review, no matter how sophisticated its conduct, duplicate this fundamental dilemma. Its aim is to ensure that similarly situated defendants are treated the same, but that no defendant is condemned to death absent full judicial consideration of the elements that render his or her situation unlike all others. The basic predicament inherent in comparative proportionality review stems from the assumption that it is possible to identify, with some legally adequate measure of exactitude, just what makes some capital cases similar but others dissimilar, and hence what renders some deathworthy but others not. Only this premise renders tenable the belief that state high courts are indeed capable, on rationally defensible grounds, of vacating the judgment of a jury whose discretionary authority, when not adequately channeled by state statutes, results in a disproportionate sentence.

On the one hand, if appellate courts are to make good on the statutory requirement that they consider both the crime and the defendant, and do so in a way that comports with the requirement of individualized sentencing, they must engage in a particularized analysis of each death sentence in order to find out what, if anything, distinguishes this case from that of others who have been sentenced to death. However, that inquiry renders it difficult if not impossible to assemble a class of

---

\(^{438}\) Tuilaepa v. California, 512 U.S. 967, 973 (1994).

\(^{439}\) Romano v. Oklahoma, 512 U.S. 1, 6 (1994).


\(^{441}\) See id. at 1145–46 (Blackmun, J., dissenting) (denying review).
defendants on the basis of which one can make a comparative judgment about whether the death penalty is or is not generally imposed and so whether any given sentence is or is not disproportionate. On the other hand, the interest in legal uniformity necessarily draws courts away from consideration of the distinguishing circumstances of any given case through the application of comparative methods—e.g., frequency analysis—that abstract from the peculiarities of individual defendants and their crimes. However, that inquiry renders it difficult if not impossible to attend adequately to the factors which, because they distinguish a case on review from those with which it is compared, may warrant sparing a defendant’s life.

The result of this dilemma, in Washington and elsewhere, is an unhappy history of decisions in which courts render judgments on the basis of considerations which, arguably, are no less arbitrary than the jury discretion its review was originally intended to remedy. “Is it,” asks Justice Handler of the New Jersey Supreme Court, “worse to kill for money or for hatred? . . . Is it worse to kill to support a gambling habit or to support a drug habit? Is it worse to kill a relative or a stranger? To pose those questions is to pose insoluble moral conundrums.” To give the appearance of solving these conundrums by conducting a proportionality review that seems to abide by the hallmarks of legal rationality, when it is based in fact on judgments of culpability, which, as a rule, are neither articulated nor defended, is to perpetuate the myth that the death penalty can indeed be administered in a principled way that comports with the claims of fairness. It is, under the guise of the law, to reproduce rather than to remedy the arbitrariness Gregg was held to resolve.

If we continue to believe that the unique nature of the death penalty requires an unusually heightened measure of due process protection, and if state high courts have failed to provide such protection, and if our three decade experiment with comparative proportionality review is a symptom of as well a contributor to that failure, then the conclusion to be drawn seems inescapable. As Justice Handler wrote, in dissenting from the New Jersey Supreme Court’s decision in State v. Martini: Today’s decision serves as further confirmation of the failure of our experiment with capital punishment. The Court’s early belief that it could fashion a constitutionally-legitimate process

Conduct of Comparative Proportionality Review

for imposing the death penalty . . . has foundered on yet another rock—proportionality review. The inconsistency, subjectivity, and moralizing evident in today’s decision are the inevitable products of a futile endeavor: the quest to devise and to apply a standard of due-process protection commensurate with the gravity of a death sentence. . . . [T]he Court must either reject its effort to carry out capital punishment or accommodate itself to the juridical brutality of imposing death without due-process protections commensurate to its awesome finality.444

The Washington State Supreme Court was perhaps refreshingly candid when, in Pirtle, it acknowledged that “[a]t its heart, proportionality review will always be a subjective judgment as to whether a particular death sentence fairly represents the values inherent in Washington’s sentencing scheme for aggravated murder.”445 However, it is not at all clear that this confession should allay any of the concerns now daily being voiced about the fundamental fairness of the death penalty’s administration. Even those who believe that the U.S. Supreme Court was mistaken in holding that comparative proportionality review is not constitutionally required should not be deluded into thinking that what is constitutionally mandated can in fact be coherently implemented, or that its conduct can salvage a body of death penalty jurisprudence that is constitutionally infirm. Although oft-cited, Justice Marshall’s claim in Godfrey v. Georgia446 is worth recalling: “The task of eliminating arbitrariness in the infliction of capital punishment is proving to be one which our criminal justice system—and perhaps any criminal justice system—is unable to perform.”447 The history of comparative proportionality review provides additional testimony regarding the law’s inability to devise and apply a standard of due process protection that is adequate to our well-founded conviction that death is indeed different.

444. Id. at 1001 (Handler, J., dissenting) (citations omitted).
447. Id. at 440 (Marshall, J., concurring).
APPENDIX A*

DATE FILED: ______________________
(to be indicated by Clerk of Supreme Court)


REPORT OF THE TRIAL JUDGE
Aggravated First Degree Murder Case

Superior Court of ______________________ County, Washington
Cause No. ______________________
State v. ______________________

INSTRUCTIONS: Please answer each question. If you do not have sufficient information to supply an answer, please so indicate after the specific question. If sufficient space is not allowed on the questionnaire form for answer to the question, use the back of the page, indicating the number of the question which you are answering, or attach additional sheets.

If more than one defendant was convicted of aggravated first degree murder in this case, please make out a separate questionnaire for each such defendant.

The statute specifies that this report shall, within thirty (30) days after the entry of the judgment and sentence, be submitted to the Clerk of the Supreme Court, to the defendant or his or her attorney, and to the prosecuting attorney.

Conduct of Comparative Proportionality Review

(1) Information about the Defendant

(a) Name: ___________________________ Date of Birth: _______
   Last, First Middle

   Sex: M □ F □
   Marital Status: Never Married □
   Married □
   Separated □
   Divorced □
   Spouse Deceased □

   Race or ethnic origin of defendant: ________________________
   (Specify)

(b) Number and ages of defendant's children:

(c) Defendant's Father living: Yes □ No □
   If deceased, date of death: _________________
   Defendant's Mother living: Yes □ No □
   If deceased, date of death: _________________

(d) Number of children born to defendant's parents: _________________

(e) Defendant's education--check highest grade completed:

   College: □ 1 □ 2 □ 3 □ 4 □
   1 2 3 4 5 6 7 8 9 10 11 12 College: □ 1 □ 2 □ 3 □ 4 □
   1 2 3 4 □ □

   Intelligence Level: Low □ Medium □ Above Average □ High □
   IQ Score: _______

Further explanation or comment:
(f) Was a psychiatric evaluation performed:  Yes ☐  No ☐

If yes, did the evaluation indicate that the defendant was:

(i) able to distinguish right from wrong?  Yes ☐  No ☐

(ii) able to perceive the nature and quality of his or her act?  Yes ☐  No ☐

(iii) able to cooperate intelligently in his or her own defense?  Yes ☐  No ☐

(g) Please describe any character or behavior disorders found or other pertinent psychiatric or psychological information:

(h) Please describe the work record of the defendant:

(i) If the defendant has a record of prior convictions, please list:

<table>
<thead>
<tr>
<th>Offense</th>
<th>Date</th>
<th>Sentence Imposed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(j) Length of time defendant has resided in:

Washington:  ___________  County of conviction:  ___________
Conduct of Comparative Proportionality Review

(2) Information about the Trial

(a) How did the defendant plead to the charge of aggravated first degree murder?:

<table>
<thead>
<tr>
<th>Option</th>
<th>☐</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guilty</td>
<td></td>
</tr>
<tr>
<td>Not Guilty</td>
<td></td>
</tr>
<tr>
<td>Not Guilty by reason of insanity</td>
<td></td>
</tr>
</tbody>
</table>

(b) Was the defendant represented by counsel?:

<table>
<thead>
<tr>
<th>Option</th>
<th>☐</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

(c) Please indicate if there was evidence introduced or instructions given as to any defense(s) to the crime of aggravated first degree murder:

<table>
<thead>
<tr>
<th>Defense</th>
<th>Evidence</th>
<th>Instruction(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excusable Homicide</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Justifiable Homicide</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Insanity</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Duress</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Entrapment</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Alibi</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Intoxication</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Other specific defenses:</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td></td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td></td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
(d) If the defendant was charged with other offenses which were tried in the same trial, list the other offenses below and indicate whether defendant was convicted:

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(e) What aggravating circumstances, as set forth in Laws of 1981, ch. 138 § 2, were alleged against the defendant and which of these circumstances were found to have been applicable?:

<table>
<thead>
<tr>
<th>Aggravating Circumstances Alleged</th>
<th>Found Applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(f) Please provide the names of each other defendant tried jointly with this defendant, the charges filed against each other defendant, and the disposition of each charge:

<table>
<thead>
<tr>
<th>Name: __________________________</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Offenses Charged</th>
<th>Disposition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conduct of Comparative Proportionality Review

Name:  ____________________________________________

<table>
<thead>
<tr>
<th>Offenses Charged</th>
<th>Disposition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(3) Information Concerning the Special Sentencing Proceeding

(a) Date of Conviction:  ________________________________

Date special sentencing proceeding commenced:  __________

(b) Was the jury for the special sentencing proceeding composed of the
same jurors as the jury that returned the verdict to the charge of
aggravated first degree murder?  Yes ☐  No ☐

If the answer to the above question is no, please explain:

(c) Was there, in the court's opinion, credible evidence of any mitigating
circumstances as provided in Laws of 1981, ch. 138, § 7?

Yes ☐  No ☐

If yes, please describe:

(d) Was there evidence of mitigating circumstances, whether or not of a
Washington Law Review Vol. 79:775, 2004

type listed in Laws of 1981, ch. 138, § 7, not described in answer
to (3)(c) above? Yes ☐ No ☐

If yes, please describe:

e) How did the jury answer the question posed in Laws of 1981, ch. 138,
§ 6(4), that is: “Having in mind the crime of which the defendant has
been found guilty, are you convinced beyond a reasonable doubt that
there are not sufficient mitigating circumstances to merit leniency?”
Yes ☐ No ☐

f) What sentence was imposed?

(4) Information about the Victim

a) Was the victim related to the defendant by blood or marriage?
Yes ☐ No ☐

If yes, please describe the relationship: ___________________________

b) What was the victim's occupation, and was the victim an employer or
employee of the defendant?

c) Was the victim acquainted with the defendant, and if so, how well?
Conduct of Comparative Proportionality Review

(d) If the victim was a resident of Washington, please state:

Length of Washington residency: _________________________

County of residence: _________________________

Length of residency in that county: _________________________

(e) Was the victim of the same race or ethnic origin as the defendant?

Yes ☐ No ☐

If no, please state the victim's race or ethnic origin:

(f) Was the victim of the same sex as the defendant?

Yes ☐ No ☐

(g) Was the victim held hostage during the crime?

Yes ☐ No ☐

If yes, for how long: _________________________

(h) Please describe the nature and extent of any physical harm or torture inflicted upon the victim prior to death:
(i) What was the age of the victim? 

(j) What type of weapon, if any, was used in the crime?

(5) Information about the Representation of Defendant

(If more than one counsel represented the defendant, answer each question separately as to each counsel. Attach separate sheets containing answers for additional counsel.)

(a) Name of counsel: 

(b) Date on which counsel was secured: 

(c) Was counsel retained or appointed? If appointed, please state the reason therefor:

(d) How long has counsel practiced law, and what is the nature of counsel's practice?

(e) Did the same counsel serve at both the trial and the special sentencing proceeding, and if not, why not?
Conduct of Comparative Proportionality Review

(6) General Considerations

(a) Was the race or ethnic origin of the defendant, victim, or any witness an apparent factor at trial?

Yes ☐ No ☐

If yes, please explain:

(b) What percentage of the population of the county is the same race or ethnic origin as the defendant?

<table>
<thead>
<tr>
<th>Race</th>
<th>Ethnic Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 10%</td>
<td>☐</td>
</tr>
<tr>
<td>10 - 25%</td>
<td>☐</td>
</tr>
<tr>
<td>25 - 50%</td>
<td>☐</td>
</tr>
<tr>
<td>50 - 75%</td>
<td>☐</td>
</tr>
<tr>
<td>75 - 90%</td>
<td>☐</td>
</tr>
<tr>
<td>Over 90%</td>
<td>☐</td>
</tr>
</tbody>
</table>

If there appears to be any reason to answer this question with respect to a county other than the county in which the trial was held, please explain:
(c) How many persons of the defendant's or victim's race or ethnic origin were represented on the jury?

   Defendant: __________________________
   Victim: __________________________

   Further explanation or comment:

(d) Was there any evidence that persons of any particular race or ethnic origin were systematically excluded from the jury?

   Yes ☐ No ☐

   If yes, please explain:

(e) Was the sexual orientation of the defendant, victim, or any witness an apparent factor at trial?

   Yes ☐ No ☐

   If yes, please explain:
Conduct of Comparative Proportionality Review

(f) Was the jury specifically instructed to exclude race, ethnic origin, or sexual preference as an issue?

   Yes ☐  No ☐

(g) Was there extensive publicity in the community concerning this case?

   Yes ☐  No ☐

(h) Was the jury instructed to disregard such publicity?

   Yes ☐  No ☐

(i) Was the jury instructed to avoid any influence of passion, prejudice or any other arbitrary factor when considering its verdict or its findings in the special sentencing proceeding?

   Yes ☐  No ☐

(j) Please describe the nature of any evidence suggesting the necessity for instructions of the type described in 6(f) through 6(i) above which were given:
(k) General comments of the trial judge concerning the appropriateness of
the sentence, considering the crime, the defendant, and other relevant factors:

(7) Information about the Chronology of the Case

(a) Date of offense: __________________

(b) Date of arrest: __________________

(c) Date trial began: __________________

(d) Date jury returned verdict: __________________

(e) Date post-trial motions ruled on: __________________

(f) Date special sentencing proceeding began: __________________

(g) Date sentence was imposed: __________________

(h) Date this trial judge's report was completed: __________________

________________________________
TRIAL JUDGE
SETTLING SIGNIFICANT CASES

Jeffrey R. Seul

Abstract: Negotiation, mediation, and other consensus-based alternatives to litigation are most often studied and defended in the context of ordinary disputes, in which liability and distributive issues are contested, but the background norms that govern the outcome of a lawsuit are not. Many consider adjudication to be the only acceptable process for addressing “significant cases”: disputes about abortion, school prayer, the environment, and other value-laden issues in which background norms are contested. I argue that this perspective is ironic because litigation, like negotiation, entails compromise. Litigation is a lottery in which the substantive values a party seeks to defend, and which it claims are absolute, may be wholly or partially discredited by the court. Furthermore, litigation merely shifts the burden of negotiation to judges. I distinguish two types of negotiation, bargaining and moral deliberation, and argue that both should be viewed as legitimate alternatives to litigation for processing disputes involving deep moral disagreement. Deliberative dispute resolution processes present important opportunities for democratic participation, and settlements resulting from them may benefit both the parties and society in ways that litigation cannot. Even where parties are incapable of engaging in genuine moral deliberation, however, settlement for strategic reasons sometimes may be a sensible alternative for parties to a significant case, and should not invite scorn. Litigation and negotiation are complementary, mutually reinforcing social processes, and each has a legitimate role to play in our nation’s moral discourse and the evolution of social norms.

INTRODUCTION ..........................................................882
I. DIFFERENCES BETWEEN ORDINARY AND SIGNIFICANT CASES .............................................888
II. TWO IRONIES OF LITIGATION .................................893
   A. Risking it All .....................................................893
   B. Letting Others Negotiate for Us .........................896
   C. Explaining the Ironies: An Implicit Reason We Litigate .........................................................900
III. IS SETTLEMENT POSSIBLE? .................................903
   A. Strategic Settlement .............................................905
   B. Moral Deliberation ..............................................907
      1. Barriers to Settlement .....................................907

INTRODUCTION

Relatively few litigants attempt to obtain a hearing before the United States Supreme Court. Parties engaged in run-of-the-mill litigation may feel they have been wronged and that justice must be done, but most cases settle without a hearing on the merits.¹ Legal disputes that have the potential to create new law on important public policy matters—one class of what Owen Fiss dubbed “significant cases” in his oft-cited polemic Against Settlement²—are, however, different. When a dispute is about abortion, affirmative action, religion, use or preservation of the natural environment, gun control, controversial medical technologies like stem cell research, or other issues involving deep value differences, settlement is rare, and talk about the possibility of settlement may seem naïve or even reckless to some.

Consider the following example. A Colorado statute establishes an eight-foot floating buffer zone around persons entering and leaving

1. See Marc Galanter, The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts, 1 J. EMPIRICAL LEGAL STUD. (forthcoming Nov. 2004) (charting a precipitous decline in the number of federal and state civil and criminal trials over recent decades, despite increases in the number of case filings, lawyers, and judges).
2. Owen Fiss, Against Settlement, 93 YALE L.J. 1073, 1087 (1984). According to Fiss, “significant cases” mark “the real divide” between proponents and opponents of settlement. See id. He challenged negotiation proponents to “speak to these more ‘significant’ cases, and demonstrate the propriety of settling them.” Id.
Settling Significant Cases

health care facilities. Within a 100-foot radius of any entrance to a hospital or clinic, those wishing to influence women seeking abortions—whether through picketing, spoken protest, distribution of literature, or otherwise—are prohibited from coming within eight feet of a woman without her consent. The statute was challenged by abortion opponents through litigation that reached the U.S. Supreme Court. After five years of protracted litigation, the Court ultimately upheld the statute.

It is not difficult to imagine ways this case could have been settled at an earlier stage, and on terms that at least partially respected the parties’ respective values and objectives, whatever one may think of the options and the likelihood of achieving them. For example, groups opposing abortion could have agreed not to protest outside health care facilities or to limit their protest activities in specified ways if the health care facilities agreed to provide women considering an abortion with literature that presents cautionary, or even openly critical, perspectives of abortion. Or, perhaps the parties could have jointly produced a video designed to inform women (and men) considering an abortion about the full range of perspectives on the social, moral, religious, and health-related issues attending their decision, and about the various forms of public and private support available to those who make one choice or the other. Perhaps the parties could have agreed to a program of optional counseling for women interested in meeting, separately or together, with counselors on both sides of the debate. Perhaps one of these suggestions could have been combined with a 24-hour pre-abortion waiting period requirement.

Regardless of one’s views on abortion, settlement of a case like this may seem improbable at best. Settlement terms like those outlined above also may seem odd to some—out of touch with parties’ respective self-understandings and values and the divergent meanings they attach to abortion and the buffer zone statute. Settlement on these or any other...

4. Id.
6. Id. at 734–35.
7. In fact, such programs already have been implemented through legislation in many other states. See Kate Zernike, 30 Years After Abortion Ruling, New Trends but the Old Debate, N.Y. TIMES, Jan. 20, 2003, at A1 (surveying federal and state abortion laws). Twenty-five states have a mandatory counseling requirement, mandatory waiting period, or both. Id. at A16. The U.S. Supreme Court has sanctioned pre-abortion waiting period requirements, provided they are not unduly burdensome. Planned Parenthood v. Casey, 505 U.S. 833, 887 (1992).
grounds would smack of the worst sort of moral relativism to others. Even those for whom these or other settlement possibilities have some practical appeal may fear that the judiciary’s role in the production of norms regarding divisive social issues could be undermined by attempts to settle such significant cases.8

In this Article, I argue that negotiation—including its facilitated variants, such as mediation and consensus-building processes—should be viewed as a legitimate alternative to litigation for addressing disputes involving deep moral disagreements. Fiss and a small cadre of other critics of alternatives to litigation are particularly opposed to settlement of these significant cases.9 With the exception of Carrie Menkel-Meadow10 and a small handful of others,11 however, few legal scholars

8. This is one of Fiss’s key objections to settlement of significant cases. Fiss, supra note 2, at 1087.

9. In addition to Fiss, the more notable detractors of settlement include Richard Delgado, Trina Grillo, Laura Nader, Judith Resnick, Stephen Yezell, and, at least in his earlier writings on the subject, David Luban. See generally Richard Delgado et al., Fairness and Formality: Minimizing the Risk of Prejudice in Alternative Dispute Resolution, 1985 WIS. L. REV. 1359 (arguing that members of minority groups are disadvantaged in mediation); Trina Grillo, The Mediation Alternative: Process Dangers for Women, 100 YALE L.J. 1545 (1991) (arguing that mediation disadvantages women); David Luban, Settlements and the Erosion of the Public Realm, 83 GEO. L.J. 2619 (1995) (arguing that settlement deprives the public of legal norms and that settlements are less just than litigated outcomes) [hereinafter Erosion]; Laura Nader, Controlling Processes in the Practice of Law: Hierarchy and Pacification in the Movement to Re-Form Dispute Ideology, 9 OHIO ST. J. ON DISP. RESOL. 1 (1993) (arguing that Alternative Dispute Resolution (ADR) proponents value harmony over justice); Judith Resnick, Failing Faith: Adjudicatory Procedure in Decline, 53 U. CHI. L. REV. 494 (1986) (criticizing trend toward promotion of settlement in Federal Rules of Civil Procedure); Stephen C. Yezell, The Misunderstood Consequences of Modern Civil Process, 1994 WIS. L. REV. 631 (arguing that changes in civil process, including increased promotion of alternatives to litigation, have weakened the justice system by removing cases from appellate scrutiny).

10. Through a series of articles, Menkel-Meadow, one of settlement’s most prolific and jurisprudentially inclined proponents, progressively develops the most robust argument for settlement of significant cases that has been offered to date. See, e.g., Carrie Menkel-Meadow, For and Against Settlement: Uses and Abuses of the Mandatory Settlement Conference, 33 UCLA L. REV. 485, 500–01 (1985) (arguing that it can be hard to distinguish “private” from “public” oriented disputes and that parties to important public disputes may often achieve a higher quality of justice through settlement than they could through litigation); Carrie Menkel-Meadow, The Lawyer as Consensus Builder: Ethics for a New Practice, 70 TENN. L. REV. 3 (2002) (examining new roles for lawyers as neutral third parties who can assist in the resolution of all types of disputes, including those involving significant public policy issues); Carrie Menkel-Meadow, Practicing “In the Interests of Justice” in the Twenty-First Century: Pursuing Peace as Justice 70 FORDHAM L. REV. 1761, 1763 (2002) [hereinafter Menkel-Meadow, Pursuing Peace] (arguing that lawyers and parties can serve their interests by “searching for consensus solutions to seemingly intractable public policy and legal disputes”); Carrie Menkel-Meadow, Toward Another View of Legal Negotiation: The Structure of Problem Solving, 31 UCLA L. REV. 754, 835–36 (1984) (arguing that “cases where the law must be clear, and in abortion, school busing, etc.” are “appropriately settled by total victory,”

884
but that some of these disputes “may still benefit from a problem-solving conception both before and after the decree”); Carrie Menkel-Meadow, The Trouble With the Adversary System in a Postmodern, Multicultural World, 38 WM. & MARY L. REV. 5, 35 (1996) [hereinafter Menkel-Meadow, Trouble] (suggesting that consensus-oriented approaches to dispute resolution can be used effectively “for policy deliberations”); Carrie Menkel-Meadow, When Litigation Is Not the Only Way: Consensus Building and Mediation as Public Interest Lawyerizing, 10 WASH. U. J. & POL’Y. 37, 39 (2002) [hereinafter Menkel-Meadow, Litigation] (exploring “how processes that enable the expression and ‘handling’ of conflict may serve the public interest as well as, if not better than, the simplistic Anglo-American conception of adversary justice or public interest litigation”); Carrie Menkel-Meadow, Whose Dispute Is it Anyway?: A Philosophical and Democratic Defense of Settlement (in Some Cases), 83 GEO. L.J. 2663, 2676 (1995) [hereinafter Menkel-Meadow, Whose Dispute?] (arguing that settlement can be seen as a form of “democratic expression” that enables parties to obtain more individualized forms of justice and even to “explore avenues for law reform”). Menkel-Meadow does not draw a tight distinction between ordinary and significant cases as I have defined them, no doubt in part because she believes that what I call ordinary cases sometimes present significant policy issues. See Menkel-Meadow, Whose Dispute? at 2667 n.24 (“It is not only our larger ‘structural’ lawsuits that raise important issues of public values, but even the ‘smallest’ of cases has significant public, as well as private, possibilities of value clarification.”). Her belief in the potential public significance of seemingly ordinary cases is a byproduct of her experience mediating automobile accident cases, in which “multicultural/racial issues” and questions regarding “standards of human behavior and responsibility” often surfaced. Id. In her most recent work, however, Menkel-Meadow focuses on settlement of what I refer to as significant cases more explicitly than she does in most of her earlier work. This suggests her acknowledgment that some cases present exceptional public policy issues and that her general defense of consensual dispute resolution processes applies to these cases with equal force. See, e.g., Menkel-Meadow, Pursuing Peace at 1763 (advocating the search “for consensus solutions to seemingly intractable public policy and legal disputes”).

11. Several other legal scholars, practicing lawyers, and mediators have also offered explicit justifications for settling significant cases. See, e.g., Margaret G. Farrell, Revisiting Roe v. Wade: Substance and Process in the Abortion Debate, 68 IND. L.J. 269, 330–60 (1993) (discussing potential benefits of consensus-oriented, “participatory adjudication” model for resolving the issues presented by Roe v. Wade, 410 U.S. 113 (1973)); Barbara Ashley Phillips & Anthony C. Piazza, The Role of Mediation in Public Interest Disputes, 34 HASTINGS L.J. 1231, 1236–44 (1983) (advocating the use of mediation in policy disputes); Riley M. Sinder et al., Promoting Progress: The Supreme Court’s Duty of Care, 23 OHIO N.U. L. REV. 71, 122–32 (1996) (arguing that the U.S. Supreme Court should exercise its authority in a manner that does not dictate solutions to social problems, but which encourages the development of consensus solutions to them). Earlier strains of legal scholarship that are not focused specifically on settlement of the types of significant cases I address—particularly the work of “legal process” scholars—also lend support to the notion of settling value-laden disputes. See, e.g., HENRY M. HART, JR. & ALBERT M. SACKS, THE LEGAL PROCESS: BASIC PROBLEMS IN THE MAKING AND APPLICATION OF LAW 645 (William Eskridge & Phillip Frickey eds., 1994) (encouraging use of negotiation and mediation for “disputes which are not susceptible of solution by reasoning from generally applicable criteria of decision”); Abram Chayes, The Role of the Judge in Public Law Litigation, 89 HARV. L. REV. 1281, 1298–1302 (1976) (discussing the changing role of courts in institutional reform, including their role in helping parties fashion negotiated remedies, often with the help of third parties functioning as mediators); Lon Fuller, The Forms and Limits of Adjudication, 92 HARV. L. REV. 353, 395 (1978) (arguing that adjudication is ill-suited to the resolution of disputes involving “polycentric” problems, in which the parties’ relationships have many, interrelated strands); Lon Fuller, Mediation—Its Form and Functions, 44 S. CAL. L. REV. 305, 325 (1971) [hereinafter Fuller, Mediation] (describing the
of negotiation and alternative dispute resolution have openly supported the prospect of settling significant cases. 12 To date, most settlement proponents have focused primarily on what one might call ordinary cases: disputes that typically occur against a background of relatively well-settled legal norms that are widely considered to be morally legitimate.

I believe negotiation should be viewed as a credible alternative to litigation for resolving disputes that raise important public policy questions. As explained more fully below, I use the term “negotiation” in two senses. The first is “strategic settlement”: pure, self-interested bargaining in which each party is willing to satisfy others’ interests solely as a strategy for satisfying its own interests. The second is collective moral deliberation, in which parties explicitly seek mutually recognizable moral grounds on which to justify the terms of their agreement. When moral deliberation produces a formal or informal agreement among the parties, the agreement necessarily results from some degree of change regarding one’s own perspective, others’ perspectives, or both. Although I do not suggest that we should actively encourage crass bargaining as the best approach for resolving disputes

“central quality of mediation” as “its capacity to reorient the parties toward each other, not by imposing rules on them, but by helping them to achieve a new and shared perception of their relationship”.

12. On the other hand, few legal scholars of negotiation and dispute resolution have openly opposed the idea. Robert Mnookin, Scott Peppet, and Andrew Tulumello advise that “some cases shouldn’t settle,” including cases in which “a party has a strong desire to create a lasting legal precedent” and those in which “a party’s interest in public vindication is so strong that it cannot be met without adjudication.” ROBERT H. MNOOKIN ET AL., BEYOND WINNING: NEGOTIATING TO CREATE VALUE IN DEALS AND DISPUTES 107 (2000) (emphasis in original). This prescription seems to be offered as a reminder to the parties themselves that litigation is the only means of obtaining a judicial precedent and (in their view) the most visible means of obtaining public vindication, rather than as a general critique of settlement of significant cases. See id.; see also Robert H. Mnookin, When Not to Negotiate: A Negotiation Imperialist Reflects on Appropriate Limits, 74 U. COLO. L. REV. 1077, 1082–90 (2003) (proposing a framework for deciding whether to negotiate in situations that raise questions regarding the appropriateness and effectiveness of negotiation). Alternative dispute resolution proponents Frank E. A. Sander and Stephen B. Goldberg suggest that, from the public’s perspective, “a court resolution might be preferable to a private settlement” of a case involving a significant policy question, such as a dispute that “raises a significant question of statutory or constitutional interpretation.” Frank E. A. Sander & Stephen B. Goldberg, Fitting the Forum to the Fuss: A User-Friendly Guide to Selecting an ADR Procedure, 10 NEGOTIATION J. 49, 60 (1994). For this reason, they argue that courts should not necessarily “encourage or assist settlement in such a case.” Id. But see WILLIAM L. URY ET AL., GETTING DISPUTES RESOLVED: DESIGNING SYSTEMS TO CUT THE COSTS OF CONFLICT 17 (1988) (“In at least some cases . . . rights-based court procedures are preferable, from a societal perspective, to resolution through interests-based negotiation.”).
Settling Significant Cases

involving deeply held values, I argue that both types of negotiation can play a legitimate role in the management and the eventual, just resolution of these disputes in various circumstances and make valuable contributions to democracy.

Litigation of a dispute involving deep moral disagreement has costs as well as benefits, from the perspectives of both the parties and the public. Negotiation also has costs and benefits. Even in the realm of significant cases, litigation’s ultimate value—to the parties and to society as a whole—can be assessed only by comparing litigation to other available means of responding to the dispute.\(^{13}\) The comparative costs and benefits of litigation versus settlement have been explored extensively with respect to the everyday types of disputes that crowd court dockets.\(^{14}\) Legal scholars have said less about the relative benefits of litigation and negotiation with respect to significant cases.

In Part I of this Article, I examine some key differences between ordinary and significant cases—differences that may make settlement of significant cases seem inappropriate or implausible to some. In Part II, I discuss two ironies that are little acknowledged and discussed, but are implicit in resistance to the idea of settling significant cases. In Part III, I respond to claims that it is impossible to settle significant cases because of the intractable nature of disputes involving deeply held values. In Part IV, I consider what litigants and the public stand to gain, and lose, when cases that present significant public policy questions are settled. I

13. See Jules Coleman & Charles Silver, *Justice in Settlements*, 4 SOC. PHIL. & POL’Y 102, 104–05 (1986) (noting that litigation and settlement each have costs and benefits); Fuller, *Mediation*, supra note 11, at 307 (arguing that we must develop the ability “to appraise the relative aptness, for solving a given problem, of the various competing forms of social ordering”); Menkel-Meadow, *Pursuing Peace*, supra note 10, at 1770 (“Claims about the fairness and ‘justice’ of all forms of process must be considered both for what they promise and do internally, as well as in relation to what else is available.”); Sander & Goldberg, *supra* note 12, at 60 (discussing advantages and disadvantages of various alternatives to litigation).

explore the relationship between settlement and deliberative democratic theory in Part V, arguing that settlement processes that invite collective moral deliberation can be an important form of democratic participation. In Part VI, I anticipate and respond to claims that my perspective represents a form of moral relativism. I also present a set of principles for designing consensual dispute resolution processes geared toward resolving disputes involving deep value differences. Fidelity to these principles, I argue, increases the likelihood that the outcomes of such processes will be morally sound, socially desirable, and durable.

I. DIFFERENCES BETWEEN ORDINARY AND SIGNIFICANT CASES

As indicated above, the types of cases I refer to as “ordinary cases” typically involve disputes that occur against a background of relatively well-settled legal norms that the parties accept as legitimate. For example, in a typical contract action, the parties do not question whether the plaintiff is entitled to damages if the defendant breached a contract with the plaintiff. They disagree about whether the defendant breached and, if the judge or jury concludes that it did, the amount of damages to which the plaintiff is entitled. Similarly, in the typical tort action, the parties dispute whether the defendant caused the plaintiff’s injuries and, if so, how much compensation is appropriate. They do not question whether a party responsible for another’s injuries is liable to the injured party.

Significant cases are different. They involve contested social norms, and the competing norms defended by the parties often are foundational to their respective worldviews. Fiss identifies four types of significant cases:

cases in which there are significant distributional inequalities;
those in which it is difficult to generate authoritative consent because organizations or social groups are parties or because the power to settle is vested in autonomous agents; those in which the court must continue to supervise the parties after judgment; and those in which justice needs to be done, or to put it more modestly, where there is a genuine social need for an authoritative

15. See Theodore M. Benditt, Compromising Interests and Principles, in COMPROMISE IN ETHICS, LAW, AND POLITICS 26, 32 (J. Roland Pennock & John W. Chapman eds., 1979) (distinguishing between disputes about ideals and disputes about interests). For a general taxonomy of dispute types, see generally GEORGI SIMMEL, CONFLICT (Kurt H. Wolff trans., 1955).
Settling Significant Cases

I give some attention to cases that exhibit the first three of these characteristics, but I focus on the final type of case for two reasons. First, I believe cases likely to produce an authoritative interpretation of law often (though, obviously, not always) also fall into one or more of the other categories. For instance, the early school desegregation and busing cases exhibited all four characteristics to some degree, but the fourth characteristic arguably is what made them most socially, politically, legally, and morally significant. Second, I consider Fiss’s and others’ objections to settlement to be most compelling with respect to the fourth type of case, and less has been offered by way of rebuttal to these objections.

Significant cases either present novel legal issues or, much to the contrary, arise against a well-settled background of legal norms that a segment of the population considers to be out-of-step with contemporary (or at least their preferred) social norms. In the first instance, background legal norms are extremely thin or nonexistent. Because the parties are not “bargaining in the shadow of the law” in any meaningful sense, it would be very difficult for them to resolve their dispute consensually based upon convergent expectations about how a court would resolve it. The most reliable type of information—a prior decision that addresses the disputed issues—is simply unavailable. A

16. Fiss, supra note 2, at 1087.
18. See Robert H. Mnookin & Lewis Kornhauser, Bargaining in the Shadow of the Law: The Case of Divorce, 88 YALE L.J. 950, 968–69 (1979) (arguing that disputants’ relative substantive and procedural rights influence negotiation behavior and outcomes). But see ROBERT C. ELICKSON, ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES 40–81 (1991) (reporting results of empirical study of animal trespass and boundary fence dispute settlements among farmers and ranchers in Shasta County, California, in which parties based settlements on informal norms rather than relevant legal principles, of which they were generally unaware); GERALD N. ROSENBERG, THE HOLLOW HOPE: CAN COURTS BRING ABOUT SOCIAL CHANGE? 125–26 (1991) (reporting results of empirical studies which, among other things, indicate that most people are unaware of decisions of the U.S. Supreme Court); Janet Cooper Alexander, Do the Merits Matter? A Study of Settlements in Securities Class Actions, 43 STAN. L. REV. 497, 505–68 (1991) (reporting results of empirical study of securities fraud settlements in which settlement amounts bore little or no relation to expected value of trial outcomes). While the average citizen may be unaware of decisions of the U.S. Supreme Court that affect their interests, representatives of social groups that typically influence or control litigation of significant cases (e.g., activists and public interest lawyers) no doubt are aware of them. Interestingly, Ian Ayres and Eric Talley argue that uncertain legal norms actually may create incentives that promote equitable agreements. Ian Ayres & Eric Talley, Solomonic Bargaining: Dividing a Legal Entitlement to Facilitate Coasean Trade, 104 YALE L.J. 1027, 1102 (1995).
notable example of the second type of case—one where existing law was challenged and unsettled as a result of a shift in prevailing social norms—is *Brown v. Board of Education*.\(^{19}\) *Brown I* upended segregationist legal norms established, legitimated and protected by *Plessy v. Ferguson*\(^{20}\) and a series of prior U.S. Supreme Court decisions.\(^{21}\) In this second type of significant case, a prior decision addressing the disputed issues exists, but one party opposes it and is hopeful that it can persuade the Court to alter or reverse the decision.

As Paul Campos explains, significant cases are difficult to resolve through the type of rational argumentation and decision-making that occurs when fundamental principles are not at stake and generally accepted legal norms make the outcome of litigation more predictable.\(^{22}\) Most difficult legal problems involve not only complicated empirical problems, but also problematic judgments concerning questions of moral value, and (often as a direct consequence of these other difficulties) various conceptually incommensurable definitions of what sorts of facts are said to constitute legal meaning. These latter types of disputes will tend not to be amenable to resolution through the procurement of more evidence via the workings of the dispute processing system, either because they involve conceptual disagreements about what should even count as evidence, or because they can’t usefully be thought of as involving evidentiary questions at all.\(^{23}\)

One might say that ordinary cases involve questions of justice with a small “j,” whereas significant cases involve questions of Justice writ large. In significant cases, the parties are trying to establish or buttress background legal norms, either by creating a new legal norm where none presently exists, or by subverting or reaffirming an existing legal norm that mirrors a favored social norm. Each party is pursuing justice in the

---

20. 163 U.S. 537 (1896).
22. PAUL F. CAMPOS, JURISMANIA: THE MADNESS OF AMERICAN LAW 63 (1998); see also DUNCAN KENNEDY, A CRITIQUE OF ADJUDICATION (FIN DE SIÉCLE) 42–43 (1997) (arguing that parties are incapable of reaching consensus in ideological disputes).
23. CAMPOS, supra note 22, at 63. As Campos explains, “within such a zone powerful competing considerations can be adduced for holding a variety of views. Furthermore such considerations can’t be refuted without recourse to some axiomatic ground of argument that others do not accept and that, precisely because it is axiomatic, cannot be argued for rationally.” Id. at 160 (emphasis in original).
Settling Significant Cases

larger sense of a legal order that affirms a social order one considers normative. The parties are pursuing their respective universalizing projects by attempting, through adjudication, to establish, alter or defend a legal rule.24

When we litigate significant cases, we do so—and say we do so—to advance or defend a claimed right or state of affairs that we believe follows axiomatically from a deeply held value, such as the sanctity of life, personal freedom, or equality.25 Though this point is obvious enough, here is a small sampling of quotes that illustrate it:

On abortion:

“[I]f you step back and consider this, that the richest people that have ever lived on the face of this earth have somehow engaged in killing one of every three of their own offspring, you have to think something is bad, bad wrong there.”26

“It really comes down to whether women will have an equal place at life’s table, whether we value children enough that we want them to be planned and wanted and cared for.”27

On the Boy Scouts’ exclusion of homosexuals:

“The Boy Scouts of America, as a private organization, must have the right to establish its own standards of membership if it is to continue to instill the values of the Scout Oath and Law in boys.”28

“The dissenters strongly embrace a sensitive and fair understanding of gay equality . . . and clearly have the better of

25. See Benditt, supra note 15, at 34 (noting that disputants’ favored policies flow from their moral convictions).
27. Id. (quoting Gloria Feldt, president, Planned Parenthood) (emphasis added).
the argument.”

In the eyes of these disputants, social justice will not exist on a grand scale unless a cherished value is actualized through enforcement of a claimed right or the institution or eradication of a particular state of social affairs. Parties to significant cases seek to establish or defend structures they consider essential to a just social order.

Deeply held values are among the raw material from which an individual’s sense of self and the identity of significant social groups to which one belongs are constructed. Significant cases often pit one moral community and at least one of its identity-defining norms against other communities and at least one of their respective, identity-defining norms. For example, in the anti-segregation phase of the civil rights struggle, African-Americans and other anti-segregationists opposed segregation laws and policies based on their commitment to the dignity and equality of African-Americans, and segregationists attempted to preserve those same laws and policies, at least in part, on the basis of their commitment to notions of federalism. The values and norms at stake in significant cases also partially define social boundaries and generate and maintain social structures and institutions in which individual and group identities and perceptions of self-interest are constructed.

It is not merely the subject matter of the dispute that gives rise to a significant case. Some disputes about school prayer generate significant cases, and others do not. The parents of students attending a private, Christian school might agree that their children should pray together, but they may disagree about how frequently the students should pray. Nor are significant cases necessarily a byproduct of disputes between members of social groups whose boundaries are partially defined by

30. Support for this proposition is readily found within the literature on social psychology. See, e.g., Kimberly A. Wade-Benzoni et al., Barriers to Resolution in Ideologically Based Negotiations: The Role of Values and Institutions, 27 ACAD. MGMT. REV. 41, 44–45 (2002) (discussing the relationship between values and individual and group identity); see also Benditt, supra note 15, at 31 (arguing, from a philosophical perspective, that “our principles often define us, at least in part”).
32. See MICHAEL WALZER, SPHERES OF JUSTICE: A DEFENSE OF PLURALISM AND EQUALITY 8 (1983) (“Men and women take on concrete identities because of the way they conceive and create, and then possess and employ social goods.”); Wade-Benzoni, supra note 30, at 43–47.
Settling Significant Cases

divergent, deeply held values. A simple contract dispute between a gun control advocate and a member of the National Rifle Association is unlikely to produce a significant case. These examples raise distributive questions that are comparatively easy to resolve, because the parties’ disagreement arises against a background of mutually accepted social or legal norms.

II. TWO IRONIES OF LITIGATION

The ardor with which significant cases are litigated is understandable in light of the depth of the litigants’ respective commitments to the values they seek to advance or defend, but two characteristics of adjudication as a process for achieving one’s objectives also make it somewhat ironic. Litigation is a lottery in which the substantive values a party seeks to defend, and which it claims are absolute, may be wholly or partially discredited by the court. Furthermore, litigation merely shifts the burden of negotiation to judges. In both of these ways, litigation, like negotiation, entails compromise.

A. Risking It All

Litigation always involves risk of an unfavorable ruling. It frequently produces binary, win-lose outcomes. When it does not, both parties lose to some extent.\(^\text{33}\) As Martin Shapiro explains, “lurking within such judicial institutions as money damages and equitable discretion are major elements of compromise.”\(^\text{34}\) Compromise is built into the system—not negotiated compromise, but compromise in the sense of choosing to forgo alternatives and accept the risk that courts will endorse a norm that is not wholly consistent with one’s perspective.\(^\text{35}\)

---

\(^{33}\) Judith Resnick, a prominent critic of alternative dispute resolution, observes that judicial decisions often are more complex than some negotiation proponents suggest. Resnick, supra note 9, at 537 n.205. Interestingly, some critics of settlement miss this point. Jules Coleman and Charles Silver, for example, argues that settlements, unlike judgments, almost never do justice because neither party prevails completely over the other. Coleman & Silver, supra note 13, at 104.

\(^{34}\) Martin Shapiro, *Compromise and Litigation*, in COMPROMISE IN ETHICS, LAW, AND POLITICS 163, 173–74 (J. Roland Pennock & John W. Chapman eds., 1979). Shapiro associates winner-take-all outcomes primarily with suits at law, where the remedy is money damages, and more integrative outcomes (i.e., a balance of the equities) with suits at equity. *Id.* at 167–68. He cites the school desegregation cases as the prime example of the latter. *Id.* at 170.

\(^{35}\) See Arthur Kuflik, *Morality and Compromise*, in COMPROMISE IN ETHICS, LAW, AND POLITICS 38, 40 (J. Roland Pennock & John W. Chapman eds., 1979) (“[T]he mere fact that the
Most disputants who litigate a significant case undoubtedly are painfully aware that the outcome of their lawsuit may be an unfavorable legal precedent, as Jules Lobel explains. Lobel, a law professor who litigated and lost many significant cases opposing the United States’ interventions in Central America during the 1980s, still questions his efforts, even as he defends them. Lobel ultimately concludes that litigation is an acceptable strategy for pursuing social change because it catalyzes political engagement and influences public discourse, contributing to the creation of a “culture of legal struggle that continually informs and inspires future generations to challenge oppressive practices.”

Whether or not one accepts Lobel’s vision of justice as endless struggle, one can appreciate litigation’s potential value in rallying others, including future generations, around one’s cause. As Lobel acknowledges, however, it is not the only potential way to mobilize others. Given litigation’s high costs, it seems odd that he makes “no special claim for . . . [its] strategic usefulness . . . as compared with other forms of political and social action.” Short of life-threatening forms of protest (e.g., hunger strikes) and acts of civil disobedience involving risk of physical retaliation or imprisonment, litigating a case all the way to the Supreme Court in an unfavorable social and political climate surely is one of the most costly signals of the strength of one’s convictions and one of the riskiest approaches for achieving one’s ultimate ends. Nonetheless, Lobel encourages lawyer-activists to “reject . . . the solution of minor improvement in favor of total redemption.” By doing so, they accept that their efforts may not improve, and perhaps may

contending parties have agreed to submit their dispute to the determination of a certain procedure will be sufficient to ascribe a compromise to them, however they fare in the end.”); id. at 53 (stating that litigants “affirm that they are prepared to make concessions to one another if, in the considered judgment of a competent judge, that is what they ought to do”) (emphasis in original); David Luban, The Quality of Justice, 66 DEN. U. L. REV. 381, 389 (1989) [hereinafter Quality] (noting that a “trial is in effect a lottery”); Shapiro, supra note 34, at 168–69 (arguing that, even in suits at law that produce winner-take-all outcomes, parties compromise by agreeing to abide by the court’s decision).

37. Id. at 1337.
38. Id. at 1353.
39. Id. at 1355.
40. Id. at 1338.
worsen, the lot of those they represent.\textsuperscript{41} Things may get worse, or stay bad much longer, before they get better.\textsuperscript{42}

The losers will likely view their loss as a temporary setback or, like Lobel, a marginal gain in light of the perceived benefits realized in terms of promoting one’s cause. Yet, legal norms articulated by courts often further entrench social norms that the losers opposed. As Riley Sinder and colleagues argue, until \textit{Brown I}\textsuperscript{43} overturned \textit{Plessy v. Ferguson},\textsuperscript{44} “segregationists could use federal and state governments to block problem-solvers from providing adequate education for blacks.”\textsuperscript{45}

Whether the costs of unsuccessful litigation are ultimately justified is a

\textsuperscript{41} It is interesting to note that many of the lawyers litigating significant cases are not among the class of persons they seek to protect. For example, Lobel, a citizen and resident of the United States, sought to protect Central Americans through litigation in U.S. courts. \textit{Id.} at 1333–34. Albion Tourgeé, who represented the losing parties in \textit{Plessy v. Ferguson}, was white. (For an account of Tourgeé’s life and legal work, see generally OTTO H. OLSEN, CARPETBAGGER’S CRUSADE: THE LIFE OF ALBION WINEGAR TOURGEÉ (1965)). Lobel’s heavy focus on the lawyer’s motivations and justifications for litigating, as well as his many florid references to activist litigators as “prophets,” visionary “fools,” “artists,” “poets,” and the like, arguably gives the lawyers’ narratives and professional identities precedence over the developing narratives, and practical circumstances, of those they represent. Lobel, \textit{supra} note 36, at 1331, 1341. To borrow a phrase from Menkel-Meadow, one cannot help but ask, “[w]ho’s dispute is it anyway?” Menkel-Meadow, \textit{Whose Dispute?}, \textit{supra} note 10, at 2663. I wonder whether those whose interests Lobel and his colleagues represented fully appreciated the negative consequences of litigation pursued and lost, as compared to other forms of political engagement. See also Derrick Bell, \textit{Serving Two Masters: Integration Ideals and Client Interests in School Desegregation Litigation}, 85 \textit{Yale L.J.} 470, 471 (1976) (questioning whether civil rights lawyers’ pursuit of their own litigation goals compromised their clients’ educational goals); Farrell, \textit{supra} note 11, at 334 (“Ideologically motivated attorneys, many of whom work with organizations with predetermined litigation programs, may consciously or unconsciously subordinate the interests of class members or subgroups of class members to their political agendas.”). As long as there are disagreements on matters involving deeply held values, the pursuit of opposing views of justice will indeed produce social struggle, but many lawyers may be too quick to file a complaint and too convinced of the necessity and inevitability of litigation to final judgment. As Lobel himself argues, “in a nation like ours, where the idea of justice historically has been attached to courts and judicial proceedings, it is inevitable that lawyers who are connected with radical social movements will introduce their struggles into the judicial arena even when they recognize that their chances of success are small.” Lobel, \textit{supra} note 36, at 1355.

\textsuperscript{42} Jack Greenberg, former director-counsel of the NAACP Legal Defense Fund (and another white civil rights litigator), advises against such high-stakes litigation when social and political conditions make a favorable ruling very unlikely. Jack Greenberg, \textit{Litigation for Social Change: Methods, Limits and Role in Democracy}, 29 \textit{Rec. Ass’n Bar City N.Y.} 320, 349 (1974) (“Lawyers ought to try to avoid creating a new \textit{Plessy v. Ferguson} and should apply energies where they will be most productive.”).

\textsuperscript{43} \textit{Brown I}, 347 US 483, 495 (1954) (overturning “separate but equal” doctrine in public education).

\textsuperscript{44} \textit{Plessy v. Ferguson}, 163 U.S. 537, 551–52 (1896) (holding that Louisiana statute requiring separate but equal railway accommodations for white and colored persons was not unconstitutional).

\textsuperscript{45} Sinder, \textit{supra} note 11, at 102.
complex (and perhaps unanswerable) question, as Lobel suggests, but there can be no doubt that there are costs.46

B. Letting Others Negotiate for Us

Litigating a significant case to final judgment is ironic in another sense. Those who litigate let others negotiate for them. Though their most cherished values are at stake, the litigants in a significant case relinquish control of its resolution to a small group of strangers. These strangers have authority to leave the litigants’ concerns unsatisfied or to produce a “balanced” outcome that is different than the balance the litigants, who have greater knowledge of their own interests, preferences and moral convictions, might otherwise strike for themselves. In effect, they let judges and juries do their compromising for them—in this case, compromise in the sense of a negotiated resolution of the dispute.

Others negotiate for us at all levels within the judicial system. Jury deliberations determine the outcome of many trials, but even bench trials are resolved by negotiation in the sense that trial court judges render their decisions through a conversation with other judges that occurs through published opinions which address the same or analogous issues. Panels of judges decide appeals. I focus on the U.S. Supreme Court in the remainder of this Part because it is the “negotiation delegate” of last resort in any significant case.47

Although negotiation among the justices of the U.S. Supreme Court is highly stylized, there is no question that they negotiate. They do not engage in the type of coarse horse-trading that sometimes occurs within the other branches of government; rather, they “accommodate” their own ideological perspectives to others’ perspectives as necessary to substantially achieve their own objectives.48 The author of an opinion

46. Some have argued that the cost of the U.S. Supreme Court’s decision in the Dred Scott case was no less than all the lives lost in the Civil War. See, e.g., EDGAR BODENHEIMER, JURISPRUDENCE: THE PHILOSOPHY AND METHOD OF THE LAW 371–72 (1962) (stating that the Civil War might have been avoided if the Court’s decision had in some way acknowledged the depth of antislavery sentiment in parts of the country); LAURENCE H. TRIBE, GOD SAVE THIS HONORABLE COURT: HOW THE CHOICE OF SUPREME COURT JUSTICES SHAPES OUR HISTORY 98 (1985) (stating that the Court’s decision “made the Civil War all but inevitable”).
47. U.S. CONST. art. III § 1.
48. See H.W. PERRY, JR., DECIDING TO DECIDE: AGENDA SETTING IN THE UNITED STATES SUPREME COURT 144–45 (1991). Although the focus of Perry’s study is the Court’s decision-making behavior on petitions for certiorari, he frequently compares the Court’s behavior at the certiorari stage to its decision-making behavior on the merits of a case. See generally id. Of course, a U.S. Supreme Court justice’s votes naturally reflect his or her “attitudes, values, or personal policy
may circulate many drafts in an effort to produce a majority, sometimes engaging in strategic behavior such as exposing less than the full court to a particular draft in an interim effort to address the requirements of a particular justice or group of justices. 49 Because a single justice cannot dictate the Court’s decision, most decisions are the product of some degree of compromise among those in the majority. Through discussion and successive draft opinions, justices tinker their way toward a conclusion that produces an outcome and supporting rationale with which each member of the majority is, by definition, at least minimally satisfied. 50

As a result of these efforts to accommodate the perspectives of other members of the Court, many judicially created norms represent an amalgam of reasons and values. 51 The U.S. Supreme Court’s holding in Roe v. Wade52 is one notable example of such a compromise result. 53 The litigants themselves might have reached this compromise if they had attempted to author a draft bill on the issue and lobbied for its passage. Pro-life activists surely do not consider the Court’s holding to be a “win-win” outcome, but at least some of them would admit that the Court attempted to balance respect for a pregnant woman’s autonomy and respect for the developing human fetus.

Unanimous decisions frequently involve a high degree of mutual accommodation. 54 The following quote from an unidentified justice, recorded by H.W. Perry, Jr. during his study of the U.S. Supreme Court’s behavior when considering petitions for certiorari, illustrates the preferences.” Jeffrey A. Segal & Albert D. Cover, Ideological Values and the Votes of U.S. Supreme Court Justices, 83 AM. POL. SCI. REV. 557, 557 (1989) (reporting results of empirical analysis demonstrating correlation between U.S. Supreme Court justices’ personal ideological beliefs as expressed in personal writing, speeches, and positions taken in the Court’s opinions).

49. See PERRY, supra note 48, at 144–45.

50. See id. at 195 (“[J]ustices . . . [bargain] when it comes to the reasoning in an opinion, if not the judgment. We know that bargaining (or accommodation) takes place at that stage.”).

51. See generally Aleksander Peczenk, Cumulation and Compromises of Reasons in the Law, in COMPROMISE IN ETHICS, LAW, AND POLITICS 176 (J. Roland Pennock & John W. Chapman eds., 1979) (arguing that compromise and aggregation of reasons is much more common in the creation of legal norms than mutually exclusive choices among them).

52. 410 U.S. 113 (1973).

53. Id. at 164–65 (recognizing a woman’s right to have an abortion prior to the point at which the fetus would be independently viable, and a state’s right to proscribe abortion thereafter, except when abortion is necessary to protect the mother’s life or health).

54. See PERRY, supra note 48, at 148; cf. SANDRA DAY O’CONNOR, THE MAJESTY OF THE LAW: REFLECTIONS OF A SUPREME COURT JUSTICE 121 (2003) (“We should never lose sight of how regrettable it is when the Court cannot find its way to agreement.”).
Chief Justice Warren was credited a lot for having a unanimous Court in Brown. The cost was having “all deliberate speed” come in. I think it would have been better to have the dissent spelled out . . . have the dissenters tell their problems, and then have a strong opinion answer the dissent rather than coming down with a weak opinion so that everyone would sign. I think it is better to acknowledge what argument there is on a controversial issue like that.55

As this remark makes clear, the justices of the U.S. Supreme Court not only negotiate, they sometimes join opinions that seem incrementalist to some signatories and too progressive to others.

On occasion, the Court likely reaches unanimity relatively easily because the justices’ perspectives on an issue already are well aligned. Conversely, the Court may sometimes reach unanimity as a result of a significant change in the perspective of at least one justice. Although it is impossible to know, the Court’s decision in Virginia v. Black,56 which upheld (in substantial part) Virginia’s statute prohibiting cross burning with an intent to intimidate,57 may be the most recent example of this. During oral argument, Justice Thomas, who ordinarily says little or nothing, interjected with several forceful comments and questions that reportedly had a visible emotional impact on other members of the Court, leading one observer to speculate that he had changed the perspectives of the other justices, who previously had expressed doubt about the constitutionality of the statute.58

Despite the pressure to produce unanimous decisions, the Court very often is divided. Sometimes it is so divided that “the” Court’s opinion consists of a thin majority perspective and multiple dissents. Other times it is so vehemently divided that we get the oddest of all opinions: a plurality opinion, in which one can barely discern the common thread that joins those viewpoints which together constitute the Court’s

55. See PERRY, supra note 48, at 148.
57. In a plurality opinion, the Court upheld the portion of the Virginia statute that outlawed burning a cross with intent to intimidate. Id. at 363. It struck down the portion of the statute that said, in effect, that burning a cross is prima facie evidence of intimidation. Id. at 367. Justice Thomas dissented with respect to the latter aspect of the plurality opinion. Id. at 388–400 (Thomas, J., dissenting).
Settling Significant Cases

decision. The Court’s opinion in *Bush v. Gore* is just one highly visible, recent example.

Whether the Court is unanimous or bitterly divided, it arrives at decisions through processes akin to those by which settling litigants, legislatures, and the public at large arrive at decisions. When the Court’s decision is unanimous, we have an example of consensual decision-making—the very result litigants achieve when they resolve their dispute without a final decision on the merits. When the Court’s decision is not unanimous, we have an example of decision-making by majority vote. This is the method by which many negotiations are resolved, from three friends’ selection of a restaurant to the U.S. Congress’s adoption of the Civil Rights Act. Even when the Court reaches its decision by putting the matter to a vote among the justices, however, some degree of “accommodation” likely has occurred among the members of the majority.

Some argue that the U.S. Supreme Court’s structure, formal and informal norms and procedures, and relative insulation from political influences eliminate crass bargaining in the process by which it arrives at its decisions, thereby encouraging moral deliberation among the justices. The Court’s mandate and institutional design, the life tenure of its members, and other factors probably do conspire to produce a higher quality of dialogue than is typical within most other institutions and spheres of social life. Ultimately, however, the Court renders its decisions on contentious matters through processes available to other institutions and groups—negotiation and voting. While bargaining may


60. See, e.g., *Alexander Bickel, The Least Dangerous Branch: The Supreme Court at the Bar of Politics* 25–26 (1962) (“Judges have, or should have, the leisure, the training, and the insulation to follow the ways of the scholar in pursuing the ends of government.”); *Ronald Dworkin, Law’s Empire* 375 (1986) (arguing that courts are more likely to “reach sound decisions about minority rights” than legislatures, because they are insulated from political pressure); *Ronald Dworkin, A Matter of Principle* 9–71 (1985) (“Judicial review ensures that the most fundamental issues of political morality will finally be set out and debated as issues of principle and not political power alone, a transformation that cannot succeed, in any case not fully, within the legislature itself.”).

61. Amy Gutmann and Dennis Thompson note that claims about the Court’s superior ability to engage in moral deliberation are empirically unfounded. Amy Gutmann & Dennis Thompson, *Democracy and Disagreement* 45 (1996). On the other hand, the Supreme Court law clerks interviewed by Perry provide some evidence that the justices sometimes do engage in bargaining. While they generally reported that the justices bargain very little with respect to their decisions on motions for certiorari, they reported that the justices bargain more in decisions on the merits, though less than the circuit judges for whom the clerks had served. See *Perry, supra* note 48, at 150–60. Numerous other studies have produced evidence of bargaining and strategic behavior among
be more prevalent in other political contexts, moral deliberation does occur, and there is no reason why it cannot be considered normative.

Viewed longitudinally, the Court appears as just another participant—albeit one with a privileged role and special powers—in a never-ending process by which we negotiate social norms. The extent to which its role in the construction and maintenance of legal norms influences the ongoing construction and maintenance of social norms is discussed below. Suffice it to say here, however, that the Court’s ultimate influence in the construction and maintenance of social norms is open to question. At most, perhaps it acts sometimes as an accelerating force, and sometimes as a crude brake, with respect to the development of dominant social norms.

C. Explaining the Ironies: An Implicit Reason We Litigate

If the explicit reason disputants litigate is to advance or defend some cherished value that conflicts with a value held by one’s adversary, why do they submit their dispute to a process in which others will negotiate a resolution that may leave them wholly or partially disappointed?

I believe that a significant, implicit reason we litigate must be that we also value the relatively pacific resolution—or at least processing—of disputes. Litigation provides a forum in which we can safely confront our adversaries and be heard—by the court itself, if not by the other party or parties. The parties’ choice among available means for addressing their dispute suggests they implicitly recognize that they, and the court before which they have brought their dispute, are participants in a larger social system that, overall, is worth maintaining and attempting to enhance. The disputants, like the rest of us, inhabit a social context in which many of their other cherished values are aligned, even though those that are the subject of their current dispute are not.


62. See infra Part IV.A.2.

63. Some would argue that at least some litigants embrace litigation as their strongest available strategy for opposition in light of the state’s monopoly on violence. History is, however, replete with violent uprisings and other forms of protest, so the decision to litigate clearly seems a decision to participate in the system—to rail against some aspect of it, rather than attempt to subvert it completely. But see ELLICKSON, supra note 18, at 144 (arguing that self-help remedies and other laws that enable parties to use force demonstrate that the state does not monopolize violence).
Settling Significant Cases

The fact that people litigate over their most deeply held values calls into question their claims that those values are absolute. This in no way implies that participants in an ideologically based dispute have an easy time choosing between litigation and its alternatives, as if litigation were just one tactic among others for expressing and attempting to advance one's moral convictions. It seems little acknowledged, however, that litigation essentially permits us to negotiate what we allege is non-negotiable, effectively creating a market through which seemingly incommensurable values can be traded. In the realm of significant cases, litigation prevents—or at least buffers us against the worst potential effects of—failures in the market for social norms. The parties’ competing values may be incommensurable in the abstract, but their actual disputing behavior seems to demonstrate that they are willing to compromise their values to some extent because they wish to inhabit a social system that is capable of containing and processing their dispute in a reasonably pacific manner.

Litigation also plays a significant role in the construction and maintenance of individual and group identities. As a form of social activism, litigation has the advantage of seemingly saving disputants

---

64. It does appear, however, that negotiated alternatives may indeed be more easily chosen in ideologically based disputes when litigation is an unattractive alternative. See Ann E. Tenbrunsel et al., The Reality and Myth of Sacred Issues in Ideologically-Based Negotiations (August 7, 2002) (unpublished manuscript, on file with author) (presenting empirical evidence that participants in disputes about deeply held values are more inclined to negotiate when they expect to lose a lawsuit).

65. Cass Sunstein argues that competing values truly are incommensurable, but that “the legal system often must put problems of incommensurability to one side, leaving those problems for ethics rather than law.” Cass R. Sunstein, Incommensurability and Valuation in Law, 92 Mich. L. Rev. 779, 820 (1994). In other words, courts sometimes must make binary choices between competing values when the parties cannot synthesize them through some unitary medium of exchange (e.g., money). If one takes a more global perspective on the role of courts in resolving value-laden disputes, however, it is clear that litigation—not “law,” but the legal process—facilitates an exchange that the parties are unwilling to make without an intermediary. In effect, the parties synthesize, or at least coordinate, their values procedurally. Their shared desire for a relatively pacific resolution of their dispute serves as a common currency; they trade the substantive values they defend (in the sense of placing them at risk) in exchange for the relative social stability they enjoy by submitting their dispute to litigation, rather than resorting to violence or taking other, more drastic action.


from making accommodative exchanges on matters they consider sacred and incommensurable, thereby sparing competing individuals and groups from the uncomfortable process of identity reconstruction that might be occasioned by a consensual resolution of the dispute. Public defense of one’s values through litigation sends powerful signals to one’s cohorts. It rallies others around the same cause, making social groups to which one belongs, and which contribute to one’s self sense, more cohesive and distinct. Litigating a case through trial and multiple appeals is a relatively public and taxing way to demonstrate the genuineness and strength of one’s convictions. An activist who attempts to advance his or her cause incrementally, through forms of political engagement that require others’ acquiescence or active support, may be perceived as making taboo trade-offs. Negotiation may seriously undermine one’s social identity or standing in the eyes of others. For some, litigation may seem to be the only means to defend values one considers non-negotiable without undermining one’s sense of self and risking sanction by one’s peers.

In the realm of significant cases, litigation permits disputants to maintain a sense of absolute conviction to a cherished value while also

---

68. This may be the answer to Malcolm Feeley’s question about why so many people resort to the court despite its limited ability to cause broad social change. Malcolm M. Feeley, Hollow Hopes, Flypaper, and Metaphors, 17 LAW & SOC. INQUIRY 745, 756–57 (1993) (reviewing GERALD N. ROSENBERG, THE HOLLOW HOPE: CAN COURTS BRING ABOUT SOCIAL CHANGE? (1991)).


70. It is easier to mobilize groups with strong collective identities than to mobilize loosely connected individuals. See Michael W. McCann, Law and Political Struggles for Social Change: Puzzles, Paradoxes, and Promises in Future Research, in LEVERAGING THE LAW: USING THE COURTS TO ACHIEVE SOCIAL CHANGE 319, 336 (David A. Schultz ed., 1998).

71. On the other hand, litigation is less costly than violent confrontation, so it arguably makes moral outrage easier to express. See Philip E. Tetlock et al., The Psychology of the Unthinkable: Taboo Trade-Offs, Forbidden Base Rates, and Heretical Counterfactuals, 78 J. PERSONALITY & SOC. PSYCHOL. 853, 868 (2000) (suggesting that moral outrage is expressed most vehemently when the perceived violation of the moral order is egregious and outrage is not costly to express).

72. Id. at 854; Leigh L. Thompson & Richard Gonzalez, Environmental Disputes: Competition for Scarcе Resources and Clashing Values, in ENVIRONMENT, ETHICS, AND BEHAVIOR: THE PSYCHOLOGY OF ENVIRONMENTAL VALUATION AND DEGRADATION 75 (Max H. Bazerman et al. eds., 1997).

73. See Benditt, supra note 15, at 31 (“[T]o compromise on matters of principle is to risk a loss of esteem, not only on the part of others, but even on one’s own part.”); Tetlock, supra note 71, at 867 (concluding, based upon experimental evidence, that “taboo trade-offs undermine core assumptions underlying relationships that are central to our conceptions of our selves and our social world”).

74. See Tetlock, supra note 71, at 855–56 (arguing that protection of sacred values has two goals, “to convince oneself of one’s moral worthiness and . . . to shore up the external moral order”).
Settling Significant Cases

enabling them to live in a world that accommodates—and which, I believe, they implicitly want to have the capacity to accommodate—a diversity of competing perspectives. Litigants go to court to serve two values, not one. One value is defended through briefs and oral argument; the other is implicit in the choice to litigate rather than secede or use violent forms of coercion. Adjudication is a more constructive response to the reality of social pluralism than dueling and other forms of violence, and that, I believe, is why we embrace it. Each of the litigants no doubt hopes that his or her perspective will prevail, but each implicitly wishes his or her perspective to prevail within a society that is capable of peacefully resolving even its most seemingly intractable conflicts.

III. IS SETTLEMENT POSSIBLE?

There is understandable skepticism about the possibility of settling the types of disputes that give rise to significant cases. As Fiss contends, “[w]e turn to the courts because we need to, not because of some quirk in our personalities.” Before considering what the disputants and the public might gain through efforts to settle the types of disputes that give rise to significant cases, it is worth considering whether settlement is possible in the first place.

The types of disputes that give rise to significant cases admittedly are among the least tractable of all conflicts. Social psychologists have discovered a number of cognitive errors and biases that systematically distort disputants’ perceptions and judgments in any negotiation, making agreements more difficult to achieve and often less than optimal when they do occur. These barriers, some of which are discussed below in

75. See HART & SACKS, supra note 11, at 4 (“The alternative to disintegrating resort to violence is the establishment of regularized and peaceable methods of decision.”); Robert M. Ackerman, Disputing Together: Conflict Resolution and the Search for Community, 18 OHIO ST. J. ON DISP. RESOL. 27, 32 (2002) (“The disputants’ willingness to submit their dispute to adjudication by a recognized tribunal is itself an affirmation of community, far more so than the self-help remedies of the blood feud, duel, or riot.”).

76. See Fiss, supra note 2, at 1089. Fiss’s skepticism about the potential for settlement of significant cases is reinforced in his brief response to McThenia and Shaffer’s critical review of Against Settlement. Compare Andrew W. McThenia & Thomas L. Shaffer, For Reconciliation, 94 YALE L.J. 1660, 1664 (1985) (arguing that the goal of ADR processes is not efficient dispute processing, but reconciliation), with Owen M. Fiss, Out of Eden, 94 YALE L.J. 1669, 1670 (1985) (arguing that courts exist to resolve disputes in a just manner when reconciliation is not possible).

Part III.B.1, may be especially high in value-laden disputes. Researchers have observed that mainstream theories and advice about negotiation, developed primarily with reference to ordinary disputes and negotiations in which financial issues predominate, are less powerful in the context of ideologically based disputes. Despite the many barriers to settlement of the types of disputes that give rise to significant cases, two types of settlements, though rare, can and do occur. One type can be thought of as strategic settlement. Strategic settlement occurs when at least one of the parties is motivated primarily by a desire to avoid the risk of an adverse decision. This party hopes that progress can be made politically, without the burden of a recently affirmed, adverse legal norm. Perhaps this party also hopes to bring “better facts” before the court at a later date. The other type of settlement results from some degree of true perspective change that produces an informal understanding or a formal agreement regarding action the parties will (or will not) take regarding matters of common concern.

Strategic settlements result from bargaining, in which the goal is maximal satisfaction of one’s own interests through compromise. Perspective change results from collective moral deliberation. As explained further in Part V, the goal of moral deliberation is the complete or partial transformation of interests, rather than their mere aggregation. The remainder of this Part demonstrates that each type of

behavior).  
78. Social psychological research confirms that, “[t]o the extent that conflicts involve people’s core values and beliefs, people will be more emotional, less able to think in an integratively complex fashion, . . . less likely to conceive or consider tradeoffs on issues involving those core values[,] . . . more rigid and single minded in their argument content and style[,] . . . [and] more likely to defend these values at any cost.” Wade-Benzoni, supra note 30, at 44.  
79. See, e.g., Tenbrunsel, supra note 64, at 5 (observing that mainstream “[b]ehavioral negotiation theories—which focus on agreements over scarce resources, competing issues and/or situations in which the basic nature of the dispute is understood . . . are not particularly useful for examining negotiations . . . that are rooted in differences in ideological beliefs”).  
80. See, e.g., GUTMANN & THOMPSON, supra note 61, at 57–63 (comparing deliberation and bargaining).  
81. Id.  
82. Bargaining often is associated with the type of interest group dynamics that are taken for granted by liberal political theorists, and deliberation often is associated with the republican political ideal. For discussions of the distinction between interest group politics and republican forms of political action, see generally Frank Michelman, Law’s Republic, 97 YALE L.J. 1493 (1988); Cass R. Sunstein, Interest Groups in American Public Law, 38 STAN. L. REV. 29 (1985). See also JÜRGEN HABERMAS, BETWEEN FACTS AND NORMS: CONTRIBUTIONS TO A DISCOURSE THEORY OF LAW and DEMOCRACY 296–97 (William Rehg trans., 1996) (associating bargaining with the
Settling Significant Cases

settlement is possible—not just in theory, but also in practice.

A. Strategic Settlement

The agreement reached shortly before the case of *Piscataway Township Board of Education v. Taxman*\(^{83}\) was to be heard by the U.S. Supreme Court is an example of strategic settlement.\(^{84}\) Sharon Taxman, a white public high school teacher in Piscataway, New Jersey, lost her job when the school district decided to eliminate a position in her department.\(^{85}\) An African-American teacher in her department, Debra Williams, who started her job on the same day as Taxman and was equally qualified, was retained.\(^{86}\) Under New Jersey law, school workforce reductions are to be made on the basis of seniority, with less senior employees being laid off first.\(^{87}\) Because both teachers had received favorable performance reviews, the school needed another basis for making its decision about whom to release from Taxman’s department.\(^{88}\) The board had broken all past seniority ties by drawing lots.\(^{89}\) In this instance, however, it made the decision on racial grounds, releasing Taxman and retaining Williams because the School Board believed its decision would promote diversity and support the goals of affirmative action.\(^{90}\)

Taxman sued, claiming reverse discrimination.\(^{91}\) The trial and appellate courts ruled in her favor,\(^{92}\) and the U.S. Supreme Court granted

---

87. Id.
88. Id.
89. Id.
90. Id.
91. Id. at 837.
92. Id. at 851, aff’d en banc, 91 F.3d 1547 (3rd Cir. 1996).
certiorari.93 Fearing a precedent with a negative effect on affirmative action programs throughout the country, a coalition of civil rights groups persuaded the school board to settle.94 The Reverend Jesse Jackson, one of the leaders of the coalition, prevailed upon the board’s president to make Ms. Taxman a generous offer of settlement, $300,000 of which would be supplied by the coalition.95

Supporters and opponents of affirmative action alike decried the Piscataway settlement.96 Many accused both sides of selling out—of “placing the route to the Supreme Court on the open market.”97 Nonetheless, the Piscataway settlement demonstrates that significant cases can be, and sometimes are, settled on strategic grounds.

A strategic settlement of a significant case is not the type of win-win, value-maximizing outcome that proponents of legal negotiation typically promise and extol. As discussed below in Part IV.A, while strategic settlement may occasionally be preferable to continued litigation from the perspective of the parties and even the public, it is a win-win outcome only in a very limited sense. While some of the various psychological barriers to settlement may have been overcome in order to achieve it—particularly disputants’ tendency toward overconfidence about their alternatives to settlement98 and reactive devaluation of an opponent’s offers of settlement99—the fundamental psychological dynamic among the parties has not shifted from simple, polarized thinking to anything approaching the sort of integratively complex thinking that proponents of settlement hope parties can achieve. Anyone who aspires to help parties to disputes involving deep value differences

94. See Estrada, supra note 84, at 215–16 (reporting events preceding settlement).
95. Id.
97. See Hentoff, supra note 96, at 22.
Settling Significant Cases

engage in the type of discourse that may result in genuine perspective change and the emergence of creative settlement options will likely view a strategic settlement as a missed opportunity and, at most, only a minor validation of the potential of negotiation to contribute to the resolution of ideological disputes.

B. Moral Deliberation

Settlements resulting from collective moral deliberation are also possible, though academics from disciplines other than law generally have been more inclined to promote deliberative dispute resolution processes and shown greater optimism about their potential to change the perspectives of parties to ideologically based disputes.\(^{100}\) Scholars in social psychology,\(^{101}\) political science,\(^{102}\) urban planning,\(^{103}\) and other fields have made important theoretical, empirical, and prescriptive contributions to our understanding of disputes involving deep value differences and productive methods for resolving them.

1. Barriers to Settlement

As indicated above, social psychologists have identified various systematic biases and errors in judgment that operate as barriers to settlement and which produce suboptimal outcomes when settlements do

\(^{100}\) Skepticism about the potential for settlement of ideological disputes is not, of course, limited to legal circles. See, e.g., Carens, supra note 67, at 129 (questioning whether “integrative compromise” is possible or desirable in disputes where one party doubts the legitimacy of another party’s perspective); Martin P. Golding, \textit{The Nature of Compromise: A Preliminary Inquiry, in COMPROMISE IN ETHICS, LAW, AND POLITICS} 3, 10–11 (J. Roland Pennock & John W. Chapman eds., 1979) (“It is far from clear that conflicts that are rooted in differences of principle or ideology . . . can be terminated either by a directly negotiated compromise or by submission to a third party.”).

\(^{101}\) See, e.g., Tenbrunsel, supra note 64, at 23–24 (presenting evidence that negotiators on either side of an environmental issue tend to be more open to compromise if they expect to lose a lawsuit on the issue); Thompson & Gonzalez, supra note 72, at 84–98 (discussing psychological obstacles to the resolution of ideological disputes and offering prescriptions for overcoming them); Wade-Benzoni, supra note 30, at 43–46, 51–53 (discussing value differences as a barrier to conflict resolution and offering suggestions for further research regarding value conflict and its resolution).

\(^{102}\) See infra Part V.

\(^{103}\) See generally \textit{JOHN FORESTER, THE DELIBERATIVE PRACTITIONER: ENCOURAGING PARTICIPATORY PLANNING PROCESSES} (1999) (encouraging deliberative approach to resolution of urban planning disputes and other difficult conflicts over resource use); \textit{LAWRENCE SUSSKIND & PATRICK FIELD, DEALING WITH AN ANGRY PUBLIC: THE MUTUAL GAINS APPROACH TO RESOLVING PUBLIC DISPUTES} (1996) (encouraging deliberative approach to resolution of urban planning disputes and other difficult conflicts over resource use).
occur.104 Skilled mediators sometimes can help disputants overcome these barriers to settlement, producing changed perspectives about the dispute and the possibilities for its resolution.105 Research and practical experience demonstrate that disputants who manage to overcome win-lose thinking and other psychological obstacles to settlement often achieve more satisfactory results through cooperation, both individually and collectively, than they would through continued conflict.106

Some psychological barriers to settlement may be particularly high in ideologically based disputes. These include the tendencies to attribute more extreme and homogeneous views to one’s opponents, and even to one’s own cohorts, than they actually hold; to resist making trade-offs one considers taboo; and to apply self-serving notions of fairness.107

Research by social psychologists establishes that participants in the types of disputes that give rise to significant cases typically view themselves as reasonable and ready to cooperate, while falsely viewing others as unreasonable and unwilling to cooperate—an attribution error that has been aptly named naïve realism.108 Researchers have found that “when each group is asked to name the cause of the dispute, each will attribute the negative aspects of the conflict to the dispositions of the other party.”109 In essence, parties on all sides of disputes involving deep value differences falsely believe that others, unlike themselves, are completely intransigent.110 One or both sides tend to “greatly exaggerate the difference between their own and the other’s belief systems in a way that exacerbate[s] the conflict.”111 In disputes regarding fundamental

---

104. See THOMPSON, supra note 77, at 124–34.
107. Thompson & Gonzalez, supra note 72, at 84–94 (discussing these and other cognitive errors and biases that operate as barriers to resolution of value-laden conflicts).
109. See THOMPSON, supra note 77, at 268.
110. The same often may be true of the lawyers that represent them.
111. THOMPSON, supra note 77, at 268. Interestingly, while liberal and conservative research subjects in studies by Robinson and his colleagues considered the views of members of the opposite group to be more extreme than they actually were, liberals, “neutrals,” and even conservatives
Settling Significant Cases

values, each side tends to attribute to the other extreme attitudes they do not actually hold.\textsuperscript{112} Disputants also tend to perceive “the other side to be more uniform in their views, whereas they perceive their own views to be more varied and heterogeneous.”\textsuperscript{113} Consequently, “[t]he general principle appears to be that coercion is viewed as more effective with our enemies than with ourselves, whereas conciliation is viewed as more effective with ourselves than with our enemies.”\textsuperscript{114}

In a study of English teachers on opposite sides of the “Great Books” debate, for example, social psychologists Robert Robinson and Dacher Keltner found that traditionalists exaggerated the position of revisionist teachers, attributing to the revisionists an exaggerated degree of ideological difference and predicting a degree of oppositional behavior that they did not exhibit.\textsuperscript{115} Traditionalists and revisionists on average selected seven books in common when asked to develop a fifteen-book reading list for an introductory college literature course.\textsuperscript{116} Another example: a March 2000 poll found that eighty-three percent of Americans believe evolution should be taught in the schools, and seventy-nine percent believe creationism should be taught. In other words, a very substantial number of Americans believe students should be exposed to both perspectives.\textsuperscript{117} These and other studies suggest there is often less ideological and practical distance between opposing moral communities than individuals on each side of a dispute realize.\textsuperscript{118}

Interestingly, individual disputants tend to believe that it is not only
their enemies whose beliefs are ideologically predetermined (unlike their own views, which they see as based upon reasonable interpretations of all available evidence); who are less capable of seeing “shades of gray”; and who are less open to conciliation. As Robert J. Robinson, Dacher Keltner, Andrew Ward and Lee Ross found in their studies of pro-choice and pro-life views of partisanship in the abortion debate and of liberal and conservative views regarding the 1986 Howard Beach racial incident, partisans also believe their own cohorts are “extreme, unreasonable, and unreachable,” though somewhat less so than their opponents.119 In other words, many of us privately see ourselves as complex moral thinkers who are surrounded by ideologues both within and outside our own moral community.

As Robinson and his colleagues explain, the consequences of this belief are unfortunate: “Partisans, accordingly, are apt to underestimate the possibility of finding common ground that could provide the basis for conciliation and constructive action; as a consequence, they could be reluctant to enter into the type of frank dialogue that could reveal . . . commonalities in interests or beliefs.”120 They also are reluctant to signal “their doubts or ambivalence to their ideological peers—lest they face coolness, suspicion, criticism, or even ostracism.”121 The bitter irony is that many individuals on each side of an ideologically based dispute may believe that all parties could profit from cooperation, but few are willing to engage in dialogue because they believe the risks of proposing it are too great in light of their low expectations about the likelihood of success.

All of this is not to say, however, that the differences between parties to the types of disputes that give rise to significant cases are not real and substantial. At times, there may be very little overlap between their underlying values (save their implicit willingness to resolve the dispute through relatively peaceful means). Perhaps more commonly, disputants embrace a similar set of values in the abstract, but place dramatically inverted emphasis on them in the context of their dispute.122 Pro-choice advocates generally do not dispute the sanctity of life and the value of the family. Pro-life advocates typically do not entirely reject the

120. Id.
121. Id. at 415.
122. See Benditt, supra note 15, at 34 (noting that conflicts of principle may arise from “different weights being given to similar principles”).
principles of individual freedom and self-determination. In the abortion context, however, each privileges one set of values over the other. Nonetheless, the fact remains that disputants tend to systematically attribute more extreme and homogenous views to both one’s opponents and one’s peers than they actually hold, and this tendency creates a significant barrier to settlement.

A second psychological barrier to settlement, which was briefly introduced above, is related to this first one. To believe that one’s adversaries and one’s cohorts hold pure and highly polarized views regarding the subject matter of a dispute is to believe that both see the dispute as a zero-sum game in which cooperation is not possible. As indicated above, however, if one privately believes cooperation may be possible, one may nonetheless be reluctant to cooperate, in part because of the social signaling costs one would incur, both in terms of one’s own reputation and in terms of the threat to group cohesiveness. When an issue is viewed as sacred, any exchange related to it is likely to be considered a “taboo tradeoff” that would invite the scorn of one’s peers and compromise one’s sense of self-worth and social standing. If one is afraid even to publicly suggest that a contested issue may be somewhat morally ambiguous, it follows that one will be even more afraid to actually act as if that is the case. As Wade-Benzoni and her colleagues explain, however, “[t]he simple truth is that it is virtually impossible for people not to make trade-offs among core values, since tradeoffs are a consequence of social existence.” Common sense and empirical evidence confirm that the degree to which one holds a value to be sacred often is situational, depending to a significant extent on how much one can afford to maintain an inflexible position.

123. See supra Part II.C.
124. Tetlock, supra note 71, at 854.
126. See, e.g., Jonathan Baron & Sarah Leshner, How Serious Are Expressions of Protected Values?, J. EXPERIMENTAL PSYCHOL.: APPLIED 183 (2000) (providing empirical evidence that values proclaimed to be sacred change based on situational factors); Tenbrunsel, supra note 64, at 21–22 (reporting evidence that suggests that parties to ideological disputes who see litigation as their alternative to agreement, and who believe they are likely to lose in court, are more likely to settle).
127. As I argue above, litigation masks this reality by creating a mechanism by which such trade-offs occur without the appearance that one has treated the core value one seeks to defend as commensurable with other values (e.g., maintaining social order). The mechanism (adjudication)
Another potentially significant barrier to settlement of ideological disputes is the tendency of disputants to be self-serving—and to be completely unaware that they are being self-serving—in their perceptions and judgments regarding fairness. Disputants tend to focus on information that favors their own interests, discount information that favors others’ interests, and thus apply fairness principles in self-serving ways, claiming for themselves more of whatever (material or symbolic) resources are available for trade than an independent third party would award to each of them. In a dispute over forest use, fishing rights or pollution, for example, environmentalists and parties representing commercial interests may have difficulty agreeing on what level of resource use or degradation is fair under the circumstances.

Wade-Benzoni and her colleagues speculate that such “egocentric interpretations of fairness” may be especially pronounced in ideological disputes. Beliefs about what is fair in ideologically based conflicts emerge from moral beliefs, which are deeply ingrained, hard to change, and associated with powerful emotions. In addition, outcomes are tied to issues of high importance, so the parties believe there is much at stake. Under such circumstances it is difficult for individuals to engage in reciprocal perspective taking (i.e., trying to see the situation from the other party’s point of view), which has been shown to help mitigate egocentrism.

Egocentrism tends to be more extreme in disputes that involve a great deal of uncertainty, either because there is disagreement about what counts as evidence, because evidence is lacking, or because the standards in light of which evidence should be judged are disputed. As we have seen, the types of disputes that produce significant cases often exhibit one or more of these characteristics.

---

128. Wade-Benzoni, supra note 30, at 43–44.
129. Id.
130. Id.
131. Id. at 44.
132. Id.
Settling Significant Cases

2. **Overcoming Barriers to Settlement**

To overcome these and other psychological barriers to settlement of an ideological dispute, one or more of the disputants must achieve a measure of perspective change sufficient to make settlement appear more attractive than continued litigation. At least one of the parties must come to view the other(s) as less extreme and more reasonable; view the dispute, or at least aspects of it, as amenable to win-win resolution; view possible trade-offs as less likely to be personally destabilizing and stigmatizing, or become more willing to accept potential ostracism; or view the facts and circumstances surrounding the dispute, and the options for its resolution, less egocentrically.

Unsurprisingly, research on attitude change indicates that the stronger one’s perspective on an issue, the more resistant one is to change. However, this same body of research indicates that perspective change is possible, even where one’s perspectives are strongly held. Most people are open to influence through deliberation and respectful persuasive appeals. Activities and experiences that tend to produce perspective change over time include: sustained exposure to alternate perspectives; appeals to shared values; experiencing the cognitive dissonance that comes from recognition of kernels of truth in another’s perspective and the potential, negative extremes of one’s own position; exploring the complexity of and internal inconsistencies within one’s own perspectives and value set; and humanizing interactions with one’s opponents.

These findings are consistent with research, discussed above, indicating that most partisans afford themselves the capacity for change (even if they deny others the same potential). Skilled mediators and other types of facilitators can help parties reduce psychological barriers to settlement by creating and guiding them through processes in which activities and experiences that produce perspective change occur.

---


134. *Id.* at 130–37.

135. *Id.* Tenbrunsel and her colleagues also recommend encouraging disputants to realistically assess their alternatives to reaching agreement, a prescription that is consistent with the mainstream theories of negotiation that they consider less powerful in the context of ideological disputes. Tenbrunsel, *supra* note 64, at 25.

136. See *supra* Part III.B.1.

137. In negotiation analytic terms, they attempt to influence the parties’ perceptions of their reservation values and the Pareto frontier, creating a zone of possible agreement where none
Perspective change may indeed be difficult to achieve, but the evidence indicates that it can and does occur. Perspective change does not necessarily entail abandonment, or even deep questioning, of one’s fundamental values. However, it does entail a minimal recognition that one’s own values and those held by one’s adversaries can co-exist and, at least to some extent, be practically accommodated. For example, mediators in Colorado helped a diverse group of state officials, activists, conservative Christians, and others achieve consensus on a plan for utilizing federal funds available for HIV prevention programs. Similarly, dialogues sponsored by the Network for Life and Choice have produced collaboration among pro-life and pro-choice activists on programs designed to minimize the number of unwanted pregnancies, assist pregnant drug addicts, and promote adoption, among other initiatives. Parties to such agreements typically do not surrender their values or beliefs, but they do achieve sufficient mutual recognition and tolerance to make possible coordinated efforts to address a problem about which all are concerned. They reach “consensus not on paradigms, value systems, or belief systems but on practical options to support together.”

Even if the respective value systems of the parties in the two disputes previously existed. See, e.g., HOWARD RAIFFA, THE ART AND SCIENCE OF NEGOTIATION 139 (1982) (explaining the concept of the Pareto frontier); HOWARD RAIFFA, NEGOTIATION ANALYSIS: THE SCIENCE AND ART OF COLLABORATIVE DECISION MAKING 110–12 (1982) [hereinafter, SCIENCE AND ART] (explaining the concept of zone of possible agreement). In the context of ideological disputes, the parties’ reservation values are, or are defined by, the sacred values they seek to defend. Sacred values may be harder to maintain when one is prevented from quickly coming to moral closure on a dilemma that implicates them. See Tetlock, supra note 71, at 867 (“If this process of reaching rapid moral closure is impeded, the mental self-control necessary for preserving taboos can become more problematic. The boundaries of the unthinkable do shift over time.”). Good faith participation in deliberative exercises with one’s ideological opponents may stimulate an attitude of moral openness, rather than moral closure.


139. Michelle LeBaron & Nike Carstarphen, Finding Common Ground on Abortion, in THE CONSENSUS BUILDING HANDBOOK: A COMPREHENSIVE GUIDE TO REACHING AGREEMENT 1031, 1032, 1045 (Lawrence Susskind et al. eds., 1999).

140. Forester considers this focus on practical alternatives, as opposed to misguided efforts to produce a synthesis of the parties’ value systems, to be the true wisdom of mediation. Forester, Dealing, supra note 138, at 489.

141. LeBaron & Carstarphen, supra note 139, at 483.
just discussed remained largely unchanged, the resolutions they achieved are markedly different from the settlement produced by the parties to the *Piscataway* case. Their settlements were not tactical moves in the service of an overarching adversarial strategy, but genuine efforts to accommodate conflicting perspectives in a manner each party presumably hoped would prove to be durable and, at least to some degree, mutually beneficial. If the parties’ fundamental moral perspectives did not change, their perspectives about each other and the possibilities for cooperation on matters where their interests are aligned certainly did.

Of course, opportunities to promote perspective change may be limited if attorneys, professional activists, and others who lead or represent social groups are conditioned to believe that it is unattainable or undesirable. Wade-Benzoni and her colleagues argue that institutions represent yet another barrier to settlement of ideologically based disputes.\(^{142}\) Institutions are social structures that influence our behavior, perceptions, and judgments,\(^ {143}\) including our expectations about what is possible and socially acceptable. They include professional cultures— their processes of education and socialization, the incentive structures that permeate them, and the norms that regulate them. To the extent that members of the legal profession, as a result of legal education, incentive structures,\(^ {144}\) and other influences, are conditioned to believe that settlement of significant cases is impossible or without potential benefit to the parties or society, this represents a powerful barrier to consensual resolution of the ideologically based disputes that give rise to them. As Jules Lobel observes, “in a nation like ours, where the idea of justice historically has been attached to courts and judicial proceedings, it is inevitable that lawyers who are connected with radical social movements will introduce their struggles into the judicial arena . . . .”\(^ {145}\) Those involved in ideologically based disputes frequently turn to lawyers for guidance and assistance in obtaining redress. If lawyers fail to see, or excessively discount, the potential value of negotiation as a strategy for serving their clients’ interests, this may make settlement of ideologically

---

143. *Id.* at 47.
144. Needless to say, lawyers often have a vested interest in litigating, not just for financial gain, but for reputational gains and ideological satisfaction. See Feeley, *supra* note 68, at 758 (arguing that lawyers enhance their own social power by “promoting the belief in an extraordinarily powerful court” capable of causing sweeping social change).
based disputes all the more infrequent, because they may fail to recognize, or may counsel their clients against exploring, possibilities for consensual resolution of the dispute.\textsuperscript{146}

Of course, there is little hope for overcoming these barriers to settlement if the disputants and their representatives do not appreciate the potential benefits of settlement, both to themselves and to society more generally. Any assessment of potential benefits must, however, acknowledge the costs of settlement. I discuss these costs and benefits in the next Part.

IV. POTENTIAL COSTS AND BENEFITS OF SETTLING SIGNIFICANT CASES

One may advocate litigation of significant cases because one believes settlement is not possible or because it is better for the parties and society in very practical ways, but I find it difficult to regard the U.S. Supreme Court as an institution peculiarly capable of making moral judgments.\textsuperscript{147} Given that the Court often compromises through a stylized blend of the same decision-making processes used by other political institutions and social groups (i.e., negotiation and voting),\textsuperscript{148} it is difficult to view the Court as a body uniquely capable of dispensing justice and of articulating “our chosen ideals.”\textsuperscript{149} Indeed, each word in this brief quote from Fiss’s \textit{Against Settlement}—our, chosen, ideals—is questionable. As others have argued extensively, the judiciary is in many senses the least democratic of the three branches of government because it is unelected and does not consult the full spectrum of its constituents before acting,\textsuperscript{150} so a Supreme Court decision arguably expresses “our”

\textsuperscript{146} Menkel-Meadow argues that lawyers not only should embrace negotiation as a credible alternative to litigation for achieving just resolutions of disputes, but that lawyers are particularly well qualified to help parties reach consensus by virtue of their greater “process consciousness.” See Menkel-Meadow, \textit{Pursuing Peace}, supra note 10, at 1763.

\textsuperscript{147} The Court has sometimes acknowledged as much. See, e.g., Roe v. Wade, 410 U.S. 113, 159 (1973) (“When those trained in the respective disciplines of medicine, philosophy, and theology are unable to arrive at any consensus [in response to the difficult question of when life begins], the judiciary, at this point in the development of man’s knowledge, is not in a position to speculate as to the answer.”).

\textsuperscript{148} See supra Part III.B.

\textsuperscript{149} Fiss, supra note 2, at 1089.

\textsuperscript{150} See, e.g., ROBERT A. DAHL, HOW DEMOCRATIC IS THE AMERICAN CONSTITUTION? 152–54 (2002) (arguing that the U.S. Supreme Court’s scope of review should be limited because the judiciary is not accountable to the public through elections). The so-called “countermajoritarian
Settling Significant Cases

ideals in a very thin sense, if one grants that it expresses them at all. Parties to litigation typically represent only a small sampling of the range of perspectives on an issue. When the Court fully validates the perspective of one of the parties, others’ perspectives obviously are discounted. When it does not fully validate either party’s perspective, perhaps the most we can say is that the Court’s opinion represents the chosen ideals of the justices who joined it, as well as those who happen to agree with them after the fact. Because many majority opinions result from some degree of “accommodation” (i.e., compromise), however, one can legitimately question whether these opinions even express the signatories’ chosen ideals as purely as Fiss and others imagine.

Adjudication is but one process for managing social conflict; the judiciary is but one institution among many that assist in social coordination and the development of social norms. Accordingly, we must consider what the parties and society gain and lose by turning to the courts to settle not only ordinary cases, but also the types of disputes that give rise to significant cases. Even in the realm of significant cases, litigation’s ultimate value can be assessed only by comparison to other available means for managing the dispute.

Too little sustained attention has been given to the comparative costs and benefits of litigation versus settlement of significant cases, because, with a few notable exceptions discussed below, almost all of the literature on legal negotiation, and most of the literature on negotiation generally, focuses on ordinary disputes. As Leigh Thompson and Richard Gonzalez explain, “[t]he negotiation literature, in a sense, conveniently sidesteps the problem of values by focusing on trade-offs of interests among parties who are already in agreement on the basic nature of the dispute.”151 Settlement of a significant case, or of a dispute that could give rise to one, involves both costs and benefits for the parties and the public, and the costs and benefits differ depending upon whether settlement occurs for strategic reasons or results from changed


151. Thompson & Gonzalez, supra note 72, at 99; see also Alexander, supra note 18, at 596 (“Our willingness to accept the [economic] model as descriptively accurate may owe much to its intuitive ring of truth in our experience of paradigm lawsuits such as simple, two-party breach of contract or personal injury cases, the sorts of cases on which the models are usually based.”).
perspectives about the dispute.

In the first major subpart below, I consider the potential costs and benefits of strategic settlement. I am not suggesting that we should actively promote coarse bargaining over litigation of significant cases, but I believe we should not be surprised, or consider it an abomination, when a strategically motivated settlement does occur. In the second major subpart below, I consider the potential costs and benefits of settlements born of genuine perspective change resulting from the disputants’ participation in types of deliberative processes that can be distinguished from the sort of coarse bargaining that produces strategic settlements. While I acknowledge the limitations and potential costs of such deliberative processes, I argue that, unlike coarse bargaining, we should actively promote deliberative processes as a method for helping parties explore whether their respective values, needs and preferences, as well as the public interest, might be better served through settlement than through continued litigation.

Whether settlement results from bargaining or moral deliberation, the parties incur two primary types of costs when a significant case is settled: opportunity costs and signaling costs.\textsuperscript{152} Settlement obviously implies foregoing the opportunity to litigate and the potential benefits litigation may produce. It also sends signals to others that may damage one’s reputation or threaten the cohesion of social groups to which one belongs. Others share in these costs because legal norms are public goods and events that affect group cohesion and inter-group relations may have ripple effects throughout a culture. The potential benefits to the parties and the public are varied and, therefore, harder to generalize. As discussed below, the relative costs and benefits of settlement take on a different character, depending upon whether they are being assessed in the context of a strategic settlement or a settlement resulting from genuine moral deliberation.

A. Strategic Settlement

When one or more of the parties to a significant case is unwilling to examine their own values and commitments critically, or to explore the

\textsuperscript{152} For a brief discussion of opportunity costs in the context of legal disputes, see MNOOKIN, \textit{supra} note 12, at 226 (“For a legal dispute, pursuing litigation is typically a client’s [best alternative to a negotiated agreement]. A rational settlement process requires that a client compare the advantages and disadvantages of a possible settlement with the opportunities and risks of litigation.”). For a discussion of signaling costs, see POSNER, \textit{supra} note 69, at 18–22.
potential for cooperation on practical matters despite their conflicting values, the case is unlikely to settle unless at least one of the parties wishes to do so for strategic reasons, as did the school district and its supporters in the Piscataway case.

1. Potential Costs and Benefits to the Disputants

From the perspective of a party involved in a significant case, the benefits and costs of strategic settlement are fairly obvious. By definition, each party offers the other something in settlement that the other values more than the outcome it desires from litigation, discounted by its estimate of the probability of an unfavorable ruling. For both parties in the Piscataway case (and, presumably, for at least some of those who advised or otherwise supported them), the perceived benefits of settlement apparently exceeded the perceived costs.

The opportunity costs incurred by one or both parties to a strategic settlement include foregoing the chance to establish or reinforce a legal norm aligned with one’s values, and the ancillary benefits that flow from the norm (e.g., public vindication of one’s perspective and the bargaining endowments legal norms afford those whom they favor). They also include the intra- and interpersonal benefits that might flow from defending those values by litigating a case to final resolution, regardless of whether one wins or loses the litigation or achieves some intermediate result: certainty that one is serving one’s deepest beliefs and principles, increased status and influence within a social movement, and stronger bonds within one’s group. When all parties tacitly acknowledge that the court is more likely to rule in favor of one of the parties, these opportunity costs theoretically are borne to a greater degree by the party favored to win the lawsuit if it is not settled.

Signaling costs result from the symbolic significance of the disputants’ behavior. In our society, litigation is one prominent method by which moral communities express themselves, becoming more cohesive in the process. Litigation is an especially public and costly way of signaling the strength of one’s convictions. When the parties settled the Piscataway case, those observers on both sides of the issue who expressed shock and dismay no doubt reacted that way, at least in part, because of the ambiguous signals the settlement sent, and

153. See MNOOKIN, supra note 12, at 101–07.  
154. See POSNER, supra note 69, at 18–22.
because of the potential for those ambiguous signals to erode the cohesiveness of the parties’ respective moral communities.\textsuperscript{155} These strong, negative reactions from some stakeholders on each side of the affirmative action issue who were not party to the lawsuit were signals to those within and outside their own communities that they believe the opposing perspectives in the dispute truly are incommensurable, despite the fact that the parties to the lawsuit apparently disagreed. Through their expressions of outrage, nonparties on both sides of the affirmative action debate signaled their view that the settling parties had defected from their respective communities by treating as commensurable that which those protesting the settlement consider to be absolute.

From the perspective of a party to a strategic settlement, the decision to settle certainly may strain relations with at least some of one’s cohorts. A group’s most ardent members may interpret the settlement as a complete renunciation of the group’s values, and thus a complete defection from the group. The signals sent by a strategic settlement are, however, ambiguous. The Piscataway school board and its supporters believed they were \textit{defending} their principles in the most strategic way possible at the moment they settled, and not that they were abandoning their principles.\textsuperscript{156} Ms. Taxman perhaps believed that she already had adequately demonstrated the merits of her perspective. Because most moral communities are not completely homogeneous, at least some of each party’s cohorts are likely to interpret the settlement not as a defection from the group, but as a principled response to a complex decision problem.

Why might a party like Ms. Taxman, who was favored to win her case, decide that the benefits of settlement outweigh the costs? In very general terms, the value of the school district’s settlement offer obviously exceeded the risk-adjusted value Ms. Taxman placed on the opportunity to have her case heard by the highest court in the land, and the benefits she hoped would flow from the hearing. Given the personal and public significance of the principles at stake in the litigation, however, how could she reach this conclusion?

As previously discussed, success in litigation is never assured, so the prospect of an unfavorable or compromise ruling naturally influences a

\textsuperscript{155} See \textit{supra} note 96 and accompanying text.

\textsuperscript{156} See Abby Goodnough, \textit{Why Piscataway Decided to Avoid Spotlight}, N.Y. TIMES, Dec. 2, 1997, at B5 (reporting that Piscataway school board members elected to settle because they were persuaded the U.S. Supreme Court would use the case as an opportunity to abolish affirmative action).
party’s decision regarding settlement. In addition, the parties’ (as well as onlookers’) relative tolerance for risk may differ. Many people prefer a certain result to a speculative result with a higher utility value (i.e., they choose the certain outcome even though the probability adjusted value of the speculative outcome is greater). 157 Furthermore, even a party engaged in an ideological dispute values things other than those principles implicated in the dispute. By accepting a monetary settlement from the school district, Ms. Taxman could claim vindication of her principles, obtain a substantial sum for her material support, and avoid the expense, stress and other costs of continued litigation. Ms. Taxman may never have intended to create a lasting precedent on reverse discrimination, but only to obtain redress for the harm she felt she had suffered. 158

Opponents of affirmative action who were upset by Ms. Taxman’s decision seem ignorant of how they seemingly expected her to bear privately all or most of the expense of establishing the public good they desired. Ms. Taxman understandably may have felt she already had borne an adequate share of the costs of defending her perspective for the benefit of others similarly situated, that she had created a significant public good through her trial and appellate court victories, and that declining the school district’s settlement offer was too great a personal cost to bear in light of the risks of continued litigation.

Sophisticated parties in Ms. Taxman’s position also are no doubt aware of the possibility of repeat play. While the settlement obviously had preclusion effect, thereby prohibiting Ms. Taxman from attempting to re-litigate her claim, 159 she may have been aware that a similarly situated person might later press his or her claim to final resolution before the U.S. Supreme Court. Ms. Taxman established a line of precedents that has significant strategic value to those who share her perspective on affirmative action. Perhaps she ultimately cared less about establishing a clear, universa legal norm against reverse discrimination than about publicly defending her own professional competence. Perhaps Ms. Taxman also believed that the signals sent by


158. McCann’s research suggests that many lawsuits on contentious moral issues are filed for the same reason that many ordinary cases are filed: to extract concessions through settlement, rather than to obtain a final judgment. See McCann, supra note 70, at 340.

159. See Nottingham Partners v. Trans-Lux Corp., 925 F.2d 29, 32 (3d Cir. 1991) (holding that an agreement dismissing civil suit bars future action on same claims).
her decision to settle, while ambiguous, were unlikely to destabilize the broader community of affirmative action opponents or subject her to insufferable ridicule among her like-minded associates.

The Piscataway school district’s motivations for settling are easier to understand, regardless of whether one agrees with its decision. Both the school district and the advocacy groups that intervened on its behalf obviously felt that the risk-adjusted value of avoiding a U.S. Supreme Court ruling on the merits of the case exceeded the value of the funds paid to Ms. Taxman to settle the dispute. The school district and its allies sought to avoid a negative precedent in hope that a positive one might be obtained at a later time—or at least that a negative precedent could be postponed indefinitely.160

In light of the way the case had evolved through trial and its first appeal, and given the defendant’s and its supporters’ reasonable fears about the probable outcome of a hearing before the Supreme Court as it was then constituted, the opportunity costs of settlement for the school district and its supporters arguably were not great. No doubt their biggest concern was the potential signaling costs of settlement. Those costs seem modest considering the high cost of an unfavorable U.S. Supreme Court decision, and the school district’s and its supporters’ strong expectation of a negative decision.161 At any rate, the Piscataway case demonstrates that settlement on strategic grounds will occasionally be desirable to at least one of the parties if a dispute evolves in a particular way, and that there may be terms on which the other party will be amenable to settlement.

2. Potential Social Costs and Benefits

From the public’s perspective, the principal cost of strategic settlement is the missed opportunity to create, reinforce, or refine a legal norm. Judge-made norms arguably can contribute to the maximization of

160. Our courts are designed to make inputs into our fund of social norms only in response to requests by citizens and governmental actors for rulings on particular issues that they cannot resolve themselves. Because the courts’ inputs carry significant, if limited, weight, parties are rightfully cautious about which issues they submit to the courts for resolution, and when. As Lisa Estrada explains, “[t]he idea that interest groups come before our nation’s courts to influence or guide the litigation process is far from original. . . . As a pioneering civil rights lawyer in the middle part of this century, Thurgood Marshall . . . meticulously select[ed] and advance[ed] court cases which fit his strategy of laying a brick-by-brick foundation of court precedents that would ultimately support a school desegregation ruling in Brown v. Board of Education . . . .” Estrada, supra note 84, at 217–18 (citations omitted).

161. See Goodnough, supra note 156, at B5.
Settling Significant Cases

public welfare by, for example, establishing fundamental rights intended to protect vulnerable members of our society, helping reduce the risk of physical injuries to acceptable levels in light of the costs of prevention, and reducing waste in the use of resources (e.g., by ensuring resources reach those who value them most, even if promises must be broken in the process). The missed opportunity to develop legal norms is one of the principal costs of settlement of significant cases that Fiss and other critics cite.

From the public’s perspective, there are at least two major reasons not to lament this opportunity cost when strategic settlements occur. One is the risk of a “wrong” decision by the Court. The other is the questionable ability of a judicial decision to contribute decisively to the resolution of social conflict involving deeply held values.

In the realm of significant cases, where perspectives are so polarized, there is good reason not to be overly solicitous of the Court’s perspective on an issue. Needless to say, it is difficult to judge, from anything like a neutral perspective, whether a given legal norm on a divisive moral issue would maximize social welfare or otherwise be morally “right.” Different norms obviously suggest different visions of society. How


163. See generally WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 312 (1987) (arguing that “the rules of Anglo-American common law of torts are best explained as if designed to promote efficiency in the sense of minimizing the sum of expected damages and costs of care”).

164. The following hypothetical illustrates this concept:

[I]f a seller (S) owns a widget that S values at $90, that one buyer (B1) values at $110, and that another buyer (B2) values at $130, an efficient legal rule is one that will induce the parties to behave in such a way that B2 will get the widget at a cost of no more than $130 and S will get at least $90. A rule under which S would keep the widget would not be efficient. Nor would a rule under which B1 ended up with the widget. (However, a rule under which S sold the widget to B1 and B1 then sold the widget to B2 would be efficient.)


165. See Fiss, supra note 2, at 1087 (arguing that settlement of significant cases leaves unsatisfied some “genuine social need for an authoritative interpretation of law”); Luban, Erosion, supra note 9, at 2623 (“Adjudication may often prove superior to settlement for securing peace because the former, unlike the latter, creates rules and precedents.”).

166. See generally RICHARD A. POSNER, Are There Right Answers to Legal Questions?, in THE PROBLEMS OF JURISPRUDENCE 197 (1990). For many affirmative action opponents, the Piscataway case fell into Professor Fiss’s fourth category of significant cases—those where “justice must be done.” See Fiss, supra note 2, at 1087. Many supporters of affirmative action undoubtedly would have felt that justice had not been done if the court had ruled in Ms. Taxman’s favor, as it was widely expected to do.
much abortion is optimal as a matter of substantive social policy: none; free access to abortion procedures through full state subsidies at any point during pregnancy, regardless of financial need; or something in-between these two poles? This question is not answerable in any meaningful, non-ideological sense in the absence of a broad consensus on the value trade-offs implicit in a given policy choice. Utilitarian policy analysis can tell us something about abortion’s role in maintaining sustainable population growth in both overpopulated regions and countries where the birth rate is undesirably low, but only the most ardent utilitarian seeking to stem or promote population growth will embrace this type of analysis as the primary basis for establishment of abortion policy.167

Some who reject utilitarian policy justifications for legal norms argue in favor of deontological norms in the form of fundamental legal rights.168 A right created by a court or legislature for one party’s benefit, however, often is a right denied to another party, or a significant interest compromised.169 Any humane, democratic society must establish and protect a set of baseline liberties and opportunities, but liberties and opportunities often clash, and locating the baseline can be contentious. Because women still are denied membership in some clubs, they suffer a (less pervasive and stigmatizing) form of one type of discrimination African-Americans suffered during the Jim Crow era. If whites are not free to exclude African-Americans from dining and other private social establishments, why should men be able to exclude women? I would not deny women the right to join clubs that still exclude them, but many Americans, including a majority of the current members of the U.S. Supreme Court,170 likely believe that some men have an interest in associating among themselves that outweighs some women’s expressed interest in associating with them, and with each other, in the facilities

---

170. See N.Y. State Club Ass’n v. City of New York, 487 U.S. 1, 2–3 (1988) (upholding New York City ordinance banning discrimination in clubs with more than 400 members, but recognizing right to expressive association which entitles “distinctively private” clubs to exclude members based upon gender and other factors).
Settling Significant Cases

from which they currently are excluded. Many people still believe that freedom to golf at Augusta National Golf Club or to join another traditionally male club is not a baseline liberty that all women should enjoy.

As Carrie Menkel-Meadow emphasizes, ours is a party-driven legal system.\(^{171}\) Parties initiate lawsuits as one approach to dealing with conflict, and they ordinarily are free to terminate them. Where intense and fundamental moral disagreement exists, it seems best to let the parties determine when a judicial norm is needed—in other words, to let them decide for themselves, as citizens immediately affected by a moral dispute, or when they believe justice must be done. It seems best not only from the perspective of the parties, but also from the public’s perspective. As \textit{Plessy v. Ferguson} and other cases one could list illustrate,\(^ {172}\) U.S. Supreme Court decisions sometimes further entrench social norms that are later widely considered to be repressive, making it much more difficult to alter them, and the structures they maintain, through political action.\(^ {173}\) The opportunity to create, reinforce or refine a legal norm that a settled case presented is not lost forever, though some cases, like \textit{Piscataway}, undeniably present especially compelling facts—for better or worse, depending on one’s perspective. If there is truly a compelling need for a judge-made norm on a contested social issue, however, another case that presents the issue is likely to emerge.

One might respond by arguing both that the U.S. Supreme Court is frequently right, even if it is sometimes wrong, and that authoritative legal norms reduce the intensity and costs of social conflict, regardless of their rightness.\(^ {174}\) However, Gerald Rosenberg’s empirical study of the effects of decades of U.S. Supreme Court decision-making in the areas of civil rights, abortion and women’s rights, the environment, reapportionment, and criminal law calls such claims into question.\(^ {175}\)


\(^ {172}\) See \textit{Plessy v. Ferguson}, 163 U.S. 537, 551–52 (1896). Other notable examples include the Dred Scott case, \textit{Scott v. Sandford}, 60 U.S. 393, 421–23, 427 (1856) (holding that slaves and former slaves were not citizens) and \textit{Buck v. Bell}, 274 U.S. 200, 206–07 (1927) (upholding Virginia statute authorizing forced sterilization of detainees of mental health institutions who were found to be “afflicted with hereditary forms of insanity, imbecility, &c.”).

\(^ {173}\) \textit{Sinder}, supra note 11, at 124.

\(^ {174}\) See, e.g., Luban, supra note 9, at 2623 (“[A]djudication may often prove superior to settlement for securing peace because the former, unlike the latter, creates rules and precedents.”).

\(^ {175}\) See ROSENBERG, supra note 18, at 337 (concluding, based upon empirical research, that
Rosenberg’s research suggests that “there is little evidence of . . . courts’ causal contributions” to social change in these areas, either as a result of the direct effects of institutional change ordered by the Court or as a result of the symbolic value of the Court’s decisions.176

Whether or not one fully accepts Rosenberg’s conclusions, most would agree that the U.S. Supreme Court cannot make all citizens instantly adopt, and behave in accordance with, its majority views.177 This is unsurprising, particularly because the Court itself does not always speak univocally. Deep differences of opinion persist within the wider population after the Court presumably “resolves” a significant policy question. We may respect the Court as a political institution, but this respect does not deter us from protesting the Court’s decisions through whatever lawful means we can, including further litigation. Rosenberg’s research suggests that the Supreme Court never single-handedly causes significant social reform.178 Indeed, Rosenberg found “[w]here there is local hostility to change, court orders will be ignored. Community pressure, violence or threats of violence, and lack of market response all serve to curtail actions to implement court decisions”).

176. Id. at 8.
177. Some scholars critical of Rosenberg’s work nonetheless agree with his narrow conclusion that the court’s rulings have little independent influence on behavior, let alone cause sweeping social change. These scholars argue, however, that Rosenberg’s approach is too “top-down” (i.e., focused on the court’s role in producing social change) and positivist. Their own approaches to the study of law and social change tend to be “bottom-up” and interpretivist, focusing, through case studies and other qualitative empirical methods, on the role of law and litigation in the mobilization of grass roots support for social reform movements and the reconstitution of identities and social meaning. For an extended exchange regarding the strengths and weaknesses of Rosenberg’s and others’ methodological approaches, see generally Michael W. McCann, Causal Versus Constitutive Explanations (or, On the Difficulty of Being so Positive . . .), 21 LAW & SOC. INQUIRY 457 (1996) [hereinafter McCann, Explanations] (rejecting positive, causal explanations of the impact of judicial decisions in favor of an interpretive approach that examines the thought and behavior of specific political actors and groups in response to judicial decisions); McCann, supra note 70 (elaborating on interpretive approach, discussing various methodological difficulties, and proposing directions for future research); Michael W. McCann, Reform Litigation on Trial, 17 LAW & SOC. INQUIRY 715 (1992) [hereinafter McCann, Reform Litigation] (critiquing Rosenberg’s methodology and comparing it to interpretive approach); Gerald N. Rosenberg, Hollow Hopes and Other Aspirations: A Reply to Feeley and McCann, 17 LAW & SOC. INQUIRY 761 (1992) [hereinafter Rosenberg, Other Aspirations] (responding to methodological and other critiques); Gerald N. Rosenberg, Knowledge and Desire: Thinking About Courts and Social Change, in LEVERAGING THE LAW: USING THE COURTS TO ACHIEVE SOCIAL CHANGE 251 (David A. Schultz ed., 1998) [hereinafter Rosenberg, Knowledge and Desire] (critiquing McCann study of pay equity reform movement on theoretical and empirical grounds); Gerald N. Rosenberg, Positivism, Interpretivism, and the Study of Law, 21 LAW & SOC. INQUIRY 435 (1996) (critiquing McCann study of pay equity reform movement on theoretical and empirical grounds).
178. See ROSENBERG, supra note 18, at 338; see also Menkel-Meadow, Pursuing Peace, supra note 10, at 1762 (observing that judicial rulings do not substantially change “the underlying social,
that the majority of Americans are unaware of the Court’s decisions in significant cases.\footnote{179}

\textit{Brown I} and the second \textit{Brown v. Board of Education}\footnote{180} decision were followed by years of inaction in some parts of the country, and, ultimately, a cascade of further litigation.\footnote{181} Rosenberg argues that

[i]he courts were ineffective in producing significant social reform in civil rights in the first decade after \textit{Brown} for three key reasons . . . First, political leadership at the national, state, and local levels was arrayed against civil rights, making implementation of judicial decisions virtually impossible. Second, the culture of the South was segregationist, leaving the courts with few public supporters . . . . Third, the American court system itself was designed to lack implementation powers, to move slowly, and to be strongly tied to local concerns. The presence of these constraints made the success of litigation for significant social reform virtually impossible.\footnote{182}

According to Rosenberg, while \textit{Brown I} was reflective of growing social disapproval of segregated education, the Court’s desegregation policies had very little practical effect until Congress and the executive branch began to actively support and implement them.\footnote{183}

Rosenberg’s conclusions obviously do not suggest that the Court can play no meaningful role in social reform, and Rosenberg himself would economic or political relations between the parties . . . .”).

179. \textit{ROSENBERG, supra} note 18, at 338.
180. 349 U.S. 294 (1955) [hereinafter \textit{Brown II}].
182. See \textit{ROSENBERG, supra} note 18, at 93. Rosenberg contends that three constraints limit the Court’s ability to produce social change: the Court’s authority is limited by the Constitution and existing precedent; the Court’s dependence upon political support to effect change; and lack of effective, autonomous means for implementing its decisions. \textit{Id.} at 336–37. He argues that the Court can produce significant social reform when some or all of these constraints are overcome as a result of the existence of sufficient political and/or social support for change. \textit{Id.} at 10–36.
183. \textit{Id.} at 105–06.
not support that position. The Court may be incapable of mandating broad acceptance of norms it sanctions, but it no doubt does contribute to the process of social change, which, of course, is highly recursive. Federal and state courts, Congress and state legislatures, and federal, state, and local executives and executive agencies are constantly engaged in what can be thought of as a “meta-dialog” on significant social issues—a dialog generated and informed by cooperation and conflict among activists, community groups, and other individuals and groups.

While the U.S. Supreme Court may be incapable of instantly changing the behavior of millions, its holdings add to an array of social processes that contribute to the ongoing development of social norms. The Court invalidates a regulation; the agency responsible for its administration modifies it, but not enough to avert further litigation. Congress legislates around a Supreme Court ruling, only to find the new legislation challenged in court. Social change surrounding divisive issues occurs gradually with inputs coming from many agents and many angles, and litigation before the nation’s courts no doubt plays a very significant role. If, as Rosenberg concluded, Supreme Court rulings are never independent causes of social change, they are likely contributing factors in a larger chain of causal events. While only a minority of citizens may follow the activities of the Supreme Court, the policies, strategies and activities of those who do—activists, lawyers, educators, politicians and other social elites—are informed and influenced by the Court’s rulings. Law surely influences and informs social norms and social meaning, even when changes in law do not immediately cause broad, corresponding social change.

184. See Rosenberg, Knowledge and Desire, supra note 177, at 254 (“The Hollow Hope is often misinterpreted as arguing that courts are irrelevant to social change and that court decisions have no impact on society.”)


186. See McCann, Explanations, supra note 177, at 459–66 (arguing that the positive model of the Court’s causal contributions to social change neglects or undervalues a host of contingent aspects of human social behavior that are influenced by law and legal struggle before the Court). But see Rosenberg, Knowledge and Desire, supra note 177, at 279–80 (noting that research for his book THE HOLLOW HOPE produced no evidence that the Court’s decisions independently changed citizens’ beliefs or enabled activists to build support for their causes).

187. See generally Lawrence Lessig, The Regulation of Social Meaning, 62 U. CHI. L. REV. 943
Settling Significant Cases

Within the constrained view of the Court’s ability to catalyze social change presented by Rosenberg, however, one of the major rewards arguably sought by at least one party to a significant case—social change mandated by the Court’s decision—appears very speculative. For this reason, one may legitimately question, as does Rosenberg, whether reformers are justified in expending scarce resources on a potential contributor to social change that can be expected to produce uncertain results even if it is facially successful. From the public’s perspective, the Court’s decisions also appear less likely to dampen social conflict or reduce its costs, because parties that the courts leave disappointed frequently register their ongoing protests through further litigation.

The point is not that the Court is impotent; the point is simply that it is less potent than many parties and members of the public likely assume. “Law is created against a pre-existing set of background norms and affects those norms in complex ways,” as Eric Posner explains. “Even if people do not like a particular [state of behavioral] equilibrium [produced by prevailing norms], it is never clear that legal intervention will improve the situation . . . . [I]ncremental changes in the law may fail to change behavior or may cause massive and unpredictable cascades, frustrating efforts to use the law to fine-tune people’s actions.” Posner concludes that, “[I]egal regulation, done poorly, will produce fewer gains than communal regulation [through the ongoing development of

———

(1995) (examining reciprocal relationship between changes in law and changes in social meanings underlying social norms).

188. Feeley suggests that activist litigants and lawyers do not really believe that a victory in court will alone produce desired social change, and therefore do not litigate with that hope. Feeley, supra note 68, at 749 (suggesting reformers make exaggerated claims about the court’s efficacy). Rosenberg persuasively refutes this argument. See Rosenberg, Other Aspirations, supra note 177, at 763–64.

189. See Rosenberg, Knowledge and Desire, supra note 177, at 258, 278–80. Just as Rosenberg seeks to encourage reform-minded litigators to adopt nonadjudicative unilateral strategies (e.g., lobbying) for achieving social change when conditions for overcoming the constraints on the Court’s ability to contribute to it are absent, I seek to encourage both those who seek reform and those who oppose it to consider utilizing an alternative to adjudicative and nonadjudicative unilateral strategies of engagement, i.e., consensual approaches to dealing with social conflict.


191. Id. at 176; see also McCann, supra note 70, at 336–37 (discussing the complexity and unpredictability of changes in perspective and behavior resulting from changes in law); Laurence H. Tribe, The Curvature of Constitutional Space: What Lawyers Can Learn from Modern Physics, 103 HARV. L. REV. 1, 28 (1989) (arguing that, paradoxically, school desegregation cases created impediments to desegregation).
social norms].”
Rosenberg’s research suggests that, when the Court does contribute to social change, its “contribution . . . is akin to officially recognizing the evolving state of affairs, more like the cutting of the ribbon on a new project than its construction.” If Rosenberg is right, then it is difficult to justify a blanket preference for seeking the Supreme Court’s perspective on each matter of deep public significance at the earliest possible moment. Furthermore, if social conflict resists the parties’ most strenuous efforts to resolve it through litigation, perhaps litigation sometimes perpetuates social conflict, rather than helping to resolve it.

From the public’s perspective, a major potential benefit of strategic settlement is, in a sense, delay—delay and a winnowing of the number of cases brought before the Supreme Court for resolution. Only failed settlement efforts can tell us how urgently a legal norm is needed. With time, the issues brought before the Court may be further refined, and the perspective of one of the parties (or yet another perspective) on the issues may come to predominate, so that creation of a legal norm becomes unnecessary or a subsequent Supreme Court decision on the matter, to the extent it is consistent with the emerging consensus, may be more widely accepted. Because judge-made law is created by a very small number of public officials (i.e., the trial and appellate judges who hear cases) in response to issues framed by as few as two individuals, and because judicial decisions typically go unchecked by the other branches of government, judge-made law should, in my view, be created cautiously and sparingly.

If this argument seems blasé—indifferent to the morally charged

192. POSNER, supra note 69, at 220.
193. See ROSENBERG, supra note 18, at 338.
194. Furthermore, producing legal norms at a feverish rate would not necessarily be a good thing. As David Luban admits, “our rapidly expanding legal system cannot tolerate a greatly accelerated adjudication rate because of the confusion and bad law that would result.” Luban, supra note 9, at 2620; see also Coleman & Silver, supra note 13, at 117–18 (arguing that reaffirmation of well-established legal precedents is wasteful).
195. For arguments (from different points on the political spectrum) favoring the abolition or limitation of judicial review on the basis that the judiciary is an inherently contermajoritarian institution, see, e.g., ROBERT H. BORK, SLOUCHING TOWARD GOMORRAH: MODERN LIBERALISM AND AMERICAN DECLINE 317–30 (1996) (blaming perceived cultural decay, in part, on antidemocratic decisions of judiciary); CASS R. SUNSTEIN, ONE CASE AT A TIME: JUDICIAL MINIMALISM AT THE SUPREME COURT 24–45 (1999) (advocating judicial restraint and arguing that it promotes a deliberative conception of democracy); MARK TUSHNET, TAKING THE CONSTITUTION AWAY FROM THE COURTS 174 (1999) (advocating abolition of judicial review on grounds that it would “distribute constitutional responsibility throughout the population”).

930
nature of the types of disputes that produce significant cases and the urgent priorities and needs of the litigants and others for whom they serve as proxies—consider that both parties in the scenario I am describing prefer settlement to continued litigation. Even if an absence of controlling precedent means the parties are not bargaining in the shadow of the law, they are at least bargaining in the shadow of the adjudicative process.196 As indicated above, my perspective regarding settlement of significant cases is contingent upon the existence of adjudication as an alternative to settlement. Consider also that another party may bring the same issue before the Court at a later date. Finally, consider the potential social costs of a “wrong” decision when compared to the potential social benefits of a “right” decision that is too far ahead of an evolving public consensus. Given the relatively slow rate of attrition among the justices of the Supreme Court and the Court’s reluctance to overturn itself, decisions that affirm existing legal norms may impede social change by revalidating those norms and extending their political “half lives,” whereas decisions that disrupt existing legal norms may have little power to effect immediate social change.

Another potential cost of settlement is diminution of the skill and experience of courts and litigators in ushering disputants through the adjudicative process, particularly its advanced stages (i.e., trial and appeal).197 While this point may seem like a bad joke to anyone who believes we live in an overly litigious society, there can be no doubt that fewer trials and appeals mean fewer opportunities for lawyers to acquire and develop the skills necessary to help citizens address their grievances through litigation when necessary. A well-functioning justice system is an important public good. My arguments about the comparative benefits of settlement are not intended to suggest otherwise; indeed, they are dependent upon the existence of adjudication as a meaningful alternative to settlement.

There seem to me to be two principal—and persuasive—responses to this concern. First, I seriously doubt that strategic settlement of significant cases will occur at such a dramatic rate as to inhibit the development of litigation skills among the lawyers who handle these types of cases or judging skills within the courts that hear them.


Significant cases likely represent a small fraction of all pending litigation, and judges seldom complain about a paucity of cases on their dockets. Furthermore, significant cases that reach the U.S. Supreme Court (or have a real prospect of reaching it) are seldom settled for purely strategic reasons, and probably seldom will be in the future. Strategic settlement is most likely to occur when the case was initiated in a clearly unfavorable climate (i.e., the plaintiff’s lawyers arguably made a serious tactical error) or the climate, though ambiguous when the suit was initiated, has become clearly unfavorable (e.g., the Supreme Court recently has denied certiorari on an unfavorable appellate court ruling in another circuit). At the U.S. Supreme Court level, examples of strategic settlement like *Piscataway* are relatively rare no doubt because these conditions are rare.\(^{198}\) For these reasons, it seems unlikely that significant cases will settle for strategic reasons so frequently that lawyers and judges become incapable of trying them.

Second, even though litigation allows lawyers to hone certain types of advocacy skills, it tends to inhibit the development of problem-solving skills among the parties themselves. Citizens’ recourse to lawyers and courts often signals their lack of capacity to solve their problem constructively without the *coercive* intervention of a third party. Given a choice between development of the skills of lawyers and judges in managing the (coercive) process of adjudication, on the one hand, and development of the parties’ own (consensual) problem solving skills, on the other, I believe the latter choice is better for society as a whole.\(^{199}\) This is a key theme of the next subpart, which examines the costs and benefits of settlements resulting from deliberative processes that produce some degree of perspective change among the parties.

**B. Moral Deliberation**

When parties to a significant case are willing to engage in a process through which they examine their own values and commitments or

\(^{198}\) On the other hand, purposefully commencing litigation with the expectation of settling it strategically may be an effective approach to achieving incremental progress in an unfavorable political climate. See McCann, *Reform Litigation*, supra note 177, at 737–39 (arguing that evidence from case studies and interviews of participants in pay equity movement demonstrates that litigation played an important role in extracting concessions through negotiation with employers, even though the movement met with little success in cases that were litigated to judgment).

\(^{199}\) In fact, like Menkel-Meadow, I also believe every lawyer’s repertoire should be broadened to include collaborative problem solving skills. See Menkel-Meadow, *Litigation*, supra note 10, at 59–61; Menkel-Meadow, *Pursuing Peace*, supra note 10, at 1763–65.
Settling Significant Cases

explore the potential for cooperation on practical matters despite their conflicting values, a settlement born of perspective change may emerge. As Lon Fuller observed long ago with respect to mediation, its central quality . . . [is] . . . its capacity to reorient the parties toward each other, not by imposing rules on them, but by helping them to achieve a new and shared perception of their relationship, a perception that will redirect their attitudes and dispositions toward one another.200

Fuller’s view of mediation is highly idealized, and not all mediators and facilitators strive for the “new and shared perceptions” that Fuller sees as the principal byproduct of mediation or have the knowledge and skill to craft processes likely to produce perspective change when it is an explicit or implicit goal.201 Nonetheless, skilled neutrals are capable of structuring processes that increase the likelihood that perspective change will occur, typically over multiple sessions spanning a significant length of time. When perspective change does occur, any potential settlement discussed by the parties, will, like a strategic settlement, have both benefits and costs for the parties and society.

1. Potential Costs and Benefits to the Disputants

In an insightful article that explores the moral and democratic value of settlement generally, Carrie Menkel-Meadow catalogs many of the benefits parties may realize when they settle a case, rather than continuing to litigate.202 These potential benefits apply equally to ordinary and significant cases. They include:

1. Party autonomy with respect to both process and outcome, a feature that furthers the goals of democracy whenever settlements are the product of genuine consent;

2. A wider range of possible outcomes that may better serve the parties’ needs and preferences;

3. The ability to express “a moral commitment to equality, precision in justice, accommodation, and peaceful coexistence of conflicting interests” through compromise.203

200. Fuller, Mediation, supra note 11, at 325.


203. Id.
(4) The ability to express fidelity to non-legal norms and principles through the terms of settlement, rather than resolving the dispute solely with reference to norms that have been embodied in law;

(5) A richer and more humane opportunity for participation than one ordinarily experiences in adjudicative processes, including greater potential for cathartic, educational, and other transformative moments;

(6) A greater ability to accommodate the values, needs and preferences of multiple stakeholders; and

(7) The exchange of more and different types of information that shed new light on the dispute and facilitate problem solving.

A further potential benefit can be added to this list when the dispute settled is the type that gives rise to a significant case. The parties may experience both satisfaction and personal growth from participating in a process that they hope will contribute to a more integrative form of social change, even if they also experience some anxiety about their decision.204

As with strategic settlements, the principal costs of settlements resulting from genuine moral deliberation are opportunity costs and signaling costs, though each takes on a somewhat different character in this context. When parties settle a significant case primarily for strategic reasons, they do not intend the terms of their agreement to be a partial or complete substitute for the legal norm they hoped to establish or reaffirm through litigation. There is no mutual acknowledgement of others’ perspectives, however incomplete or imperfect that acknowledgement may be, when a significant case is settled. The parties merely call a truce in what they assume will be an ongoing social struggle.

In contrast, settlements resulting from collective moral deliberation are intended to resolve the underlying dispute or some aspect of it, or to achieve incremental progress toward the dispute’s ultimate resolution. One forgoes or postpones the chance to establish or reinforce a legal norm aligned with one’s values, and the ancillary benefits that one hopes would flow from it, in return for an agreement that represents a partial advancement or preservation of the substantive values one had hoped could be advanced or preserved more completely through litigation. One also may forgo the opportunity for more complete retribution for, or

204 See Kuflik, supra note 35, at 51 (arguing that participation in sincere dialogue is an intrinsic good).
Settling Significant Cases

public vindication of, past harms.205

The opportunity costs incurred when a settlement of a U.S. Supreme Court case results from moral deliberation are greater than those incurred through a strategic settlement, because the parties’ respective investments in cooperative (or at least non-hostile) activities is greater. When a strategic settlement occurs, the primary motivation of at least one of the parties is fear of defeat in litigation. That party intends to carry on the fight in different forums, or to support others similarly situated in their fights before the Court once the political winds have shifted or the Court has been reconstituted. When a settlement results from genuine moral deliberation, the parties are, in a sense, giving up the fight—or, rather, electing to re-channel at least some of the energy they were investing in it. The settlement, and whatever activities or constraints flow from it, becomes a primary strategy for promoting one’s values going forward.

As with strategic settlements, parties elect to incur the opportunity costs associated with terminating litigation—or, where litigation has not yet commenced, of foreclosing it as an option—because they believe the benefits of settlement outweigh these costs. A settlement based on some measure of genuine perspective change may in fact provide each party more of what it wants than would success in court, particularly in light of the risk that a favorable outcome in litigation may not produce real social change where political will is lacking. It also may create a foundation for ongoing cooperation by the parties that delivers further benefits to each over time. Litigants who do not settle may rationalize the risk of loss by viewing the possible setback as temporary—a step on the path to ultimate victory, and therefore not truly a compromise of the value they seek to defend. While their commitment to that value may remain absolute, however, the value itself has no legal force following a decisive loss in litigation. Through settlement, the loser’s values would have become at least partially actualized, and further activism and cooperation might produce further, incremental improvements—and, perhaps ultimately, conditions in which one’s cherished value becomes as fully actualized as one hoped it would through litigation.

Signaling costs also assume a different character in the context of

205. However, as Martha Minow explains in the context of political settlements in the wake of mass violence, settlements sometimes can address the desire for retribution and public vindication through remedies such as truth commissions and reparations. MARTHA MINOW, BETWEEN VENGEANCE AND FORGIVENESS: FACING HISTORY AFTER GENOCIDE AND MASS VIOLENCE 52–117 (1998).
settlements resulting from genuine moral deliberation. A settlement of a significant case that has a cooperative or value-integrating dimension is more likely to be viewed as betrayal of one’s cohorts or even a defection from the group. Furthermore, a settlement in which a significant number of members of a social group participate, or which is supported by a significant number of a group’s members, may threaten the group’s cohesion and disrupt established identities. When one defends one’s values, one also defends one’s identity, both as an individual and as a member of any group to which those values are central.\(^\text{206}\) Exposing one’s values to genuine inquiry and potential transformation may alter our current self-perceptions, others’ perceptions of us, and even others’ perceptions of themselves.

The risk that settlement presents to the stability of social groups simply underscores the extent to which most social groups are not monolithic in terms of their members’ value orientations, needs, and preferences. Litigation tends to mask divisions within interest groups, leaving internal differences unexamined and unchallenged, because legal argumentation is dialectical. Advocates representing groups in conflict typically make polarized arguments that leave little room for ambiguity and admit few exceptions (or at least do not admit the claim made by one’s opponent as an exception). In contrast, settlement processes designed to produce perspective change may destabilize groups by permitting expression of alternate viewpoints, thereby forcing the groups to confront internal diversity and inconsistencies. In ideological conflicts, competing groups frequently have both hawkish and dove-like members.\(^\text{207}\) As Robinson and his colleagues found, a substantial number of the members of an identity group likely believe that they privately hold views that are more moderate than those held by the average member of the group.\(^\text{208}\)

Social psychologist Herbert Kelman argues that the development of coalitions across conflict lines is critical to the resolution of deep-seated

---

206. For a general discussion regarding the construction and defense of individual and group identity, including the role of values, see GLYNIS M. BREAKWELL, COPING WITH THREATENED IDENTITIES 98–100 (1986) (discussing the construction and defense of individual and group identity, including the role of values).

207. See generally Herbert C. Kelman, Coalitions Across Conflict Lines: The Interplay of Conflicts Within and Between the Israeli and Palestinian Communities, in CONFLICTS BETWEEN PEOPLE AND GROUPS 236 (S. Worchem & J.A. Simpson eds., 1993) (considering ways in which constructive interactions between pro-negotiation factions within conflicting groups can influence group members who oppose negotiation and reconciliation).

Settling Significant Cases

social conflict. Peaceful social change ultimately requires the progressive transformation of relationships within identity groups, as well as relations between them. As Kelman explains, conciliatory interactions between members of opposing groups help promote intra-group change as well as inter-group change, provided those who participate in the interactions maintain sufficient credibility within their own groups. It seems reasonable to assume, for example, that moderate white citizens’ experiences interacting with blacks during the civil rights era played a role in transforming other white citizens’ perspectives on racial issues. Conflict resolution practitioners often employ informal, nonbonding dialog processes in an effort to develop coalitions between members of opposing groups, in the hope that these interactions might eventually promote such intra-group change, creating the conditions from which informal cooperation and formal agreements between groups can emerge. Indeed, it is in all groups’ interests to produce an agreement that their more extreme members can minimally support; otherwise, the agreement is unlikely to prove durable.

Public settlements—formal or informal agreements that involve some level of coordinated action or inaction—impose signaling costs on those who participate in them. As with opportunity costs, however, participants in settlements accept them because they believe the cumulative costs of foregoing litigation are outweighed by the cumulative benefits of cooperation. Those who incur these costs alter the social environment in a way that may decrease the signaling costs for those who later elect to cooperate with members of opposing groups. In other words, a settlement may sometimes provide social “cover” in the same way that a U.S. Supreme Court ruling can. By affirming the equal rights and dignity of black citizens, the Supreme Court’s landmark civil rights decisions no doubt lowered the potential reputational costs incurred by whites inclined to express non-racist perspectives among whites who continued to hold racist views, or whose views were undisclosed. As Robinson and his colleagues observe, when one sees other members of one’s group expressing perspectives widely thought to be taboo, it becomes easier to express those perspectives oneself.

Before concluding this Part, I wish to return briefly to Fiss’s principal

209. See Kelman, supra note 207, at 254.
210. Id. at 240–42.
211. Id. at 238–39.
212. See Robinson, supra note 108, at 416.
objections to settlement of significant cases. Of the four types of significant cases identified by Fiss, I believe three can be viewed as expressions of skepticism about whether mutually beneficial settlements of the fourth type of case are truly achievable. Expressed in simple terms, Fiss’s list of significant cases comprises those involving (1) significant power imbalances between the parties, (2) representational complexities, (3) enforcement complexities, or (4) differences in deeply held values.\(^\text{213}\) While each of these characteristics may exist independently of the others, I believe cases that arise from differences in deeply held values frequently exhibit one or more of the other characteristics. When this is true, the presence of one or more of the other characteristics amplifies questions about whether it is possible to do justice through settlement.\(^\text{214}\) From Fiss’s perspective, judges ensure that justice is done not only by making authoritative declarations of law, but by producing their declarations through a process that Fiss believes ensures that weaker parties are treated fairly; the interests of those affected by the outcome but not participating in the process directly are protected; and the obligations resulting from the Court’s judgment are actually performed. An authoritative declaration of law would be morally suspect, of little practical consequence, or both if any of the latter three conditions were absent.

Power imbalances, which may be partially generated by existing legal norms, will often affect the terms of a settlement. Assuming the weaker individual or group is adequately represented, however, I would not interfere with the parties’ decision to settle, even if I objected to the terms. Adequate representation may come in the form of strong and determined members of social groups that are at odds with one another, competent counsel, or a combination of the two. Strong individuals at the vanguard of a group’s cause typically are the protagonists in disputes that give rise to significant cases, and the plaintiffs (or plaintiff’s chief supporters) in any litigation that ensues. Their lawyers can be extremely

---

213. Fiss, \textit{supra} note 2, at 1087.

214. I realize that Fiss demands a separate justification for settlement of each type of case. The bulk of this article is my effort to provide that justification with respect to his fourth type of case, those that he views as requiring an authoritative interpretation of law. \textit{Id.} The present discussion is a minimal effort to justify settlement where the other three characteristics are present. As indicated above, I give less attention to these characteristics because they have received sustained attention by many others and because I consider Fiss’ objections to settlement of cases where he believes an authoritative interpretation of law is required to be his most powerful line of argument against settlement.
capable and vigorous advocates. If these people believe perspective change significant enough to create a foundation for improved social relations has occurred, why should their decision not be respected, especially because they have litigation as an alternative and those similarly situated individuals who are not participants in the settlement remain free to initiate a lawsuit if they disapprove of the settlement?

Many join Fiss in expressing concern about negotiation’s potential for disempowerment of members of minority groups, women, the poor, and members of other socially disadvantaged groups when compared to litigation. While this risk is real, it is minimized significantly in any well-facilitated deliberative process—perhaps minimized to the point that the risk of disempowerment is less than that which weaker individuals face in court. One role of the facilitator(s) of any well-managed deliberative process is to ensure that participants have equal voice and that all parties have thought critically about a proposed agreement before they consent to it.

Other scholars see the ways in which well-facilitated consensual processes actually can be empowering for those who participate in them, including members of disadvantaged groups. The participants, who typically would be heard relatively little during a trial and not at all upon appeal, are encouraged by facilitators to contribute to the dialogue. Indeed, skilled facilitators endeavor to structure and conduct deliberative


processes in ways that make it safe for participants to speak openly and honestly. Participants acquire new information about their adversaries and themselves that will inform, and may transform, their own perspectives, interests and objectives, empowering them to make better decisions about how to achieve their objectives in the dispute, whether through settlement, litigation, or some other form of social action. With litigation as an ever-present alternative to settlement, it is very difficult to see how a well-facilitated deliberative process can disempower those who participate in it.

Problems of representation and generation of authoritative consent arguably are greater in litigation than in settlement. If adequacy of representation and consent are values that Fiss and others wish our justice system to promote, then, on several levels, litigation serves them less well than consensual dispute resolution processes. In the context of significant cases, small panels of judges establish norms on highly divisive issues that are binding upon citizens generally, the vast majority of whom will have had no meaningful voice in the litigation and no opportunity to consent. Furthermore, interest groups that supposedly are represented through the parties to litigation—whether African-Americans, abortion opponents, evangelical Christians, or Libertarians—are seldom monolithic in their views. Litigants often purport to express the unified perspective of groups that, in fact, are often internally heterogeneous. As discussed above, they often frame issues in polarized ways that fail to capture the full range of perspectives on a moral problem that are held by the group’s members. Deliberative processes typically can be structured to make room for expression of a range of perspectives, including the divergent perspectives held by members of a single social group. The outcomes they produce typically do not bind those who have not consented.

Fiss’s claim that the need for ongoing court supervision and

219. See, e.g., Richard Chasin et al., From Diatribe to Dialogue on Divisive Public Issues: Approaches Drawn from Family Therapy, 13 MEDIATION Q. 323, 331–37 (1996) (discussing efforts made to prevent reenactment of the conflict in dialogue process and to promote expression of previously unexpressed feelings and perspectives); LeBaron & Carstarphen, supra note 139, at 1031 (discussing the importance of ground rules for creating a safe environment).

220. Some of the problems generated by the presence of pluralistic parties in negotiations are addressed in RAIFFA, SCIENCE AND ART, supra note 137, at 465–83.

221. Luban demonstrates that some settlements nonetheless produce negative externalities, shifting costs to parties who were unrepresented in the process. See Luban, supra note 35, at 404–05. This possibility is the main reason well-designed deliberative processes include representatives of the broadest possible range of stakeholders.
enforcement following resolution of some types of disputes cannot be addressed through settlement is also questionable. Fiss is primarily concerned with institutional reform litigation—cases that are frequently resolved by consent decrees that contemplate the Court’s continued involvement in the implementation of their terms. There is no reason why courts cannot play a role in the post-settlement relations among parties to other types of significant cases when the parties see advantage in the Court’s continued involvement. For example, had the parties in the abortion clinic buffer zone case discussed in the introduction to this Article agreed that groups opposing abortion would limit their protest activities outside health care facilities in specified ways if the health care facilities provided women considering an abortion with literature that presents cautionary perspectives, they could have asked the Court to enter their agreement as a consent decree and to monitor their compliance with it.

In sum, if parties to a significant case achieve a measure of genuine perspective change and wish to settle their case as a result, I believe that decision is entirely legitimate—provided, as explained below, that basic liberties and opportunities are protected. Social conflict is at least partially a product of conflicting values and meanings. When parties feel that previously conflicting values and meanings have been brought into sufficient harmony (or at least sufficiently constructive tension) that, on balance, there are more grounds for cooperation than for continued conflict, there is no reason to expect or encourage the parties to continue to litigate. I may personally disagree with the parties’ decision, and I may even speak or act in protest of it, but I would not deprive them of their ability to decide. Responding to Fiss’s and Luban’s opposition to settlement in defense of the undifferentiated interests of an abstract public, Menkel-Meadow asks rhetorically, “Whose dispute is it anyway?” Particular individuals and groups suffer discrete or generalized harms, and litigation is one way they may voice their grievances. Deliberative processes enable these individuals and groups


223. This, of course, is one of the contemporary roles of courts, as Chayes was first to observe. Chayes, supra note 11, at 1298–1302.

224. See infra Part V.

to choose for themselves between the state of affairs that a given settlement promises and the state of affairs that might eventually exist after obtaining a judgment in their favor.

2. Potential Social Costs and Benefits

As with strategic settlement, there are at least two principal costs, from the public’s perspective, of a settlement resulting from genuine moral deliberation. The first is the missed opportunity to create, reinforce, or refine a legal norm. The second cost is the lost opportunities for courts and lawyers to enhance their skills at processing disputes through adjudication. These costs are discussed at length above, so I will not repeat that discussion here.

These costs of settlement must be assessed in light of the costs of litigating significant cases, of which I believe one particularly stands out. Because parties currently turn to adjudication to address their moral disagreements so frequently and so quickly, our society’s capacity for problem-solving, among both citizens and their representatives (i.e., those who advocate their causes and public officials), arguably is not as robust as it could be. I believe that compulsively litigating the types of disputes that give rise to significant cases all the way to a supposedly definitive judicial “resolution” inhibits the development of our capacity to manage our toughest problems. Disputes present opportunities to build social capital. When we litigate conflicts, we undertake our most difficult conversations in a way that is overly mediated—mediated, that is, through the impersonal strictures of the judicial process, where our primary conversation partner is the Court—rather than having these conversations directly with other affected parties, or with minimal mediation among the parties, through a facilitated, non-binding negotiation process. As a result, we practice a weaker form of democracy than we otherwise could.

From this perspective, a trial of a significant case is a failure, and a

226. See supra Part IV.A.2.
227. See Sinder, supra note 11, at 124.
Settling Significant Cases

settlement resulting from genuine moral deliberation is a sign of democratic health. As Riley Sinder, John Lopker and Ronald Heifitz have argued, if we require the judiciary to dispense answers and corrective remedies where authoritative texts do not clearly supply them, the judiciary may stifle the democratic process and retard social progress.230 As they explain, “the Court’s issuance of a technical decision balancing rights and interests may merely perpetuate [social] work avoidance by establishing the next judicial battleground in which adamant perceptions of self-interest will compete.”

By contrast, the primary potential social benefits of settling a significant case through a process of genuine moral deliberation are the development of citizens’ capacity to manage their most difficult conflicts more constructively, the strengthening of social bonds, and, of course, the meaningful, substantive contributions to social change that may result from the settlement. In the next Part, I describe how the promise of collective moral deliberation may be realized in the context of settlement of significant cases.

V. SETTLEMENT AND THE PRACTICE OF DELIBERATIVE DEMOCRACY

The capacity for social problem solving is a key feature of what Benjamin Barber calls “strong democracy.”232 As he explains, “[w]here weak democracy eliminates conflict . . . , represses it . . . , or tolerates it . . . , strong democracy transforms conflict.”233 One form of weak democracy identified by Barber is what he calls “juridicial democracy,” in which social conflict is resolved “through deferring to a representative judicial elite.”234 Barber’s notion of strong democracy is an expression of what Amy Gutmann, Dennis Thompson and other political theorists call deliberative democracy.235 From a deliberative perspective, the goal of democracy is the transformation of political

230. See Sinder, supra note 11, at 124; see also Kuflik, supra note 35, at 50 (arguing that “controversy is sometimes more to be welcomed than lamented, for it can become the occasion for persons to broaden their perspectives and enlarge their understanding”).

231. See Sinder, supra note 11, at 124.


233. Id. (emphasis in original).

234. Id. at 142 (emphasis in original).

235. See GUTMANN & THOMPSON, supra note 61, at 1.
preferences through rational dialogue, rather than the mere aggregation of preferences through bargaining or voting.\(^{236}\) Transformation of preferences is synonymous with what I refer to in this Article as perspective change. Settlement processes that create opportunities for perspective change through collective moral deliberation can be important forums for democratic participation.

A key tenet of deliberative democracy is that “political choice, to be legitimate, must be the outcome of [collective] deliberation about ends [as well as means].”\(^{237}\) Theorists differ on the extent to which well-constructed deliberative processes ensure the moral integrity of the outcomes they produce. At one end of the spectrum is Jürgen Habermas, who posits an ideal discursive process from which, he argues, morally justifiable decisions necessarily follow.\(^{238}\) At the other end are Gutmann and Thompson.\(^{239}\) While they believe that deliberative processes which satisfy specified conditions are likely to produce outcomes that are morally sound rather than just politically expedient, they argue that an outcome is not morally acceptable if it denies any party certain basic liberties and opportunities.\(^{240}\) I consider the differences in these perspectives in the discussion of moral relativism that follows in Part VI.A. The key point for present purposes is that the negotiation processes imagined by deliberative democrats involve moral deliberation about practical alternatives, not brute bargaining over them.\(^{241}\)

Reason giving is the core practice of deliberative democracy. As Gutmann and Thompson explain, from a deliberative perspective, “when citizens or their representatives disagree morally, they should continue to reason together to reach mutually acceptable decisions.”\(^{242}\) Gutmann, like Barber, advises disputants to “avoid taking a moral conflict to court if possible.”\(^{243}\)


\(^{237}\) Id. at 5 (emphasis in original).

\(^{238}\) See infra note 303 and accompanying text.

\(^{239}\) See GUTMANN & THOMPSON, supra note 61, at 17–18.

\(^{240}\) Id.

\(^{241}\) The form of negotiation that leads to strategic settlement is what proponents of deliberative democracy and some legal scholars, such as Duncan Kennedy, call bargaining. The form of negotiation that produces perspective change regarding the nature of a dispute and the possibilities for its resolution is what proponents of deliberative democracy call deliberation and what Kennedy calls dialogue. See KENNEDY, supra note 22, at 43–44 (comparing dialogue and bargaining); see also GUTMANN & THOMPSON, supra note 61, at 55–63 (comparing deliberation and bargaining).

\(^{242}\) GUTMANN & THOMPSON, supra note 61, at 1.

\(^{243}\) Gutmann, supra note 169, at 1.
Settling Significant Cases

Deliberation lies somewhere between bargaining (i.e., compromise with a focus on interests) and arguing (i.e., trying to persuade others to abandon their own reasons in favor of one’s own), and involves elements of both.\footnote{Diego Gambetta, “Claro!”: An Essay in Discursive Machismo, in DELIBERATIVE DEMOCRACY 19, 19 (Jon Elster ed., 1998).} Deliberative outcomes must be supported by “reasons that are recognizably moral in form and mutually acceptable in content.”\footnote{GUTMANN & THOMPSON, supra note 61, at 57.} One party need not be converted to another’s perspective, but each party must be able to acknowledge that every other party is defending a legitimate moral value.\footnote{Theodore Benditt argues that acknowledgment of “the sincerity and earnestness of one’s opponent” is a legitimate basis for compromise. Benditt, supra note 15, at 35. In my view, it is not enough to believe that one’s opponent is sincere. One must also believe one’s opponent’s perspective is grounded in a legitimate moral perspective. Hitler no doubt sincerely believed that genocide was a justifiable means of “purifying” the German population and culture, but absolute genetic and cultural homogeneity are not morally defensible values.} Dialogue of this quality is not easy to achieve. It requires an uncommon degree of openness to others’ perspectives and demonstrated respect for those with whom one fundamentally disagrees.\footnote{Golding, supra note 100, at 17 (“Where full reciprocity exists, the parties recognize each other as moral equals despite their relative bargaining strengths.”).} All participants in a truly deliberative process must be genuinely “motivated to find reasons that can be accepted by others.”\footnote{GUTMANN & THOMPSON, supra note 61, at 53. Participation in deliberative processes admittedly takes considerable time and effort—more time and effort than many social activists will be willing to invest. See Iris Marion Young, Activist Challenges to Deliberative Democracy, 29 POL. THEORY 670, 682 (2001) (arguing that a deliberative process “co-opts the energy of citizens committed to justice, leaving little time for mobilizing people” for other activities intended to produce reform). The succession of unsuccessful legal challenges to discriminatory laws and policies over the century preceding the Brown I decision demonstrates, however, that alternative processes for achieving social change also take considerable time and effort. See, e.g., Tushnet, supra note 215, at 169–72 (reviewing history of anti-segregation litigation preceding Brown I).} One of the key characteristics that distinguishes deliberation from pure bargaining is the attitude with which participants engage in the process.\footnote{See Benditt, supra note 15, at 26–27 (differentiating bargaining from “compromise” based upon the parties’ attitudes); see also Carens, supra note 67, at 136 (“Even in moral arguments, one approaches disagreements quite differently if one is seeking areas of mutual agreement and if one is willing to be persuaded than if one merely wishes to score a few logical points and to refute the other’s case.”); Kuflik, supra note 35, at 44 (arguing that “democracy is in danger of degenerating into a generalized ‘prisoner’s dilemma’ unless compromise is premised upon “the genuine expression of mutual respect”).}

Deliberative processes need not, however, be completely free of
bargaining. To continue an example begun above, if the members of Augusta National Golf Club and women who seek membership in it were to agree, as a consequence of a genuinely deliberative process intended to avert litigation, that men and women will have exclusive use of the club on alternating Saturdays and Sundays, and that both men and women can use the club during the remainder of the week, I suspect most proponents of deliberative democracy would approve. While this hypothetical outcome could be partially motivated by the parties’ respective fears about the risks of litigation, it also could reflect their changed perspectives regarding the possibilities for accommodating the seemingly conflicting values of inclusiveness and gender equality, on the one hand, and freedom of association, on the other. Through the process, each party would offer reasons to support the outcome that the other may come to recognize as morally legitimate, even if each party is inclined to privilege one moral principle over others served by the parties’ decision. The outcome is a product of deliberation to the extent that the parties’ motivation for settlement is grounded in mutual respect and an appreciation of the moral quality of the reasons offered by others. The outcome is the product of crass bargaining to the extent that parties consider it a prudent compromise in light of the costs and risks associated with litigation.

Campos and others are highly skeptical about the power of reason to resolve moral conflict (at least when the reasons are offered by a judge or group of judges). From Campos’ perspective, a judicial decision is not reasoned unless competing moral principles have been brought into complete harmony. In a pluralistic society such as ours, where values considered legitimate by large numbers of people often are in fundamental conflict, this typically will be impossible.

Participants in deliberative processes seldom completely abandon their own moral perspectives for those of another party, nor should they be expected to do so. Even the best-designed and best-managed consensual dispute resolution processes are unlikely to produce a

---

250. Gutmann & Thompson, supra note 61, at 71–73 (arguing that the deliberative perspective accepts bargaining with respect to some issues in dispute (e.g., contested empirical questions) provided the agreement as a whole is founded on mutually recognizable moral reasons). See also Carens, supra note 67, at 133 (arguing that bargaining can play a legitimate role in democratic politics, even when it detracts from deliberation).

251. See Campos, supra note 22, at 160 (arguing that decisions embracing one moral value to the exclusion of others are “arational”).

252. Gutmann & Thompson, supra note 61, at 93 (“The aim of a [deliberative] process is not necessarily to induce citizens to change their first-order moral beliefs.”).
complete “harmonic convergence” among parties involved in a heated moral dispute. When perspective change sufficient to produce some form of agreement occurs, however, and when the reasons given by each of the participants to justify their decision have a moral quality that is recognized by each of the other parties, I believe a form of collective reasoning has occurred. Reasoning has occurred in the sense that people who wish to serve multiple values—the substantive values underlying one’s own perspective; democratic participation; self-determination; the pacific resolution of disputes; even, to some extent perhaps, the values defended by other stakeholders—collectively embrace those values, albeit with varying emphasis, as reasons for some concrete policy or action. The choice between one party’s moral perspective and another’s need not be binary for the outcome of moral deliberation to be considered reasoned and coherent. Deliberation can bridge moral perspectives.

To the extent deliberative forms of social and political engagement succeed in reorienting understandings and relationships among citizens divided by deep value differences, they may often contribute as much or more to the evolution of social norms than would a U.S. Supreme Court decision. While much of the discussion among political theorists regarding deliberative democracy focuses on promoting and enhancing moral deliberation among elected officials, some proponents of deliberative democracy also hope to increase the quantity and quality of deliberation between representatives and citizens, as well as directly among citizens. Deliberative democracy is practiced both when public

253. Deliberative democracy is not about achieving harmony at the expense of the substantive values the parties seek to advance through deliberative processes. See, e.g., GUTMANN & THOMPSON, supra note 61, at 93 (“Deliberative reasoning is not correctly represented if it is described as giving more weight to the value of mutual respect or deliberation than (for example) to the sanctity of life.”).

254. See Kuflik, supra note 35, at 50 (arguing that “a morally appropriate balance between valid concerns is not to be equated with a compromise of moral integrity”).

255. See GUTMANN & THOMPSON, supra note 61, at 42 (advocating the creation of more “deliberative forums” that bring “previously excluded voices into politics”). New York University law professor Larry Kramer considers deliberative democracy to be anti-populist because, in his view, its procedural requirements “can be met only by small bodies far removed from direct popular control.” Larry Kramer, We the People: Who Has the Last Word on the Constitution?, BOSTON REV., Feb.-Mar. 2004, at 14, 19 (2004). While it becomes more difficult to meet the procedural requirements of deliberative democracy as the size of a group increases, the requirements can nonetheless be met in groups with several hundred members. Proponents of deliberative democracy envision deliberative processes being employed in the U.S. House of Representatives, for example, or by state legislatures. See GUTMANN & THOMPSON, supra note 61, at 46–47 (encouraging deliberation among legislators). Furthermore, the proceedings of smaller groups, including groups
officials deliberate and when deliberations among citizens are purposefully structured and managed to provide inputs into broader political processes. In order to maximize the potential for settlements of significant cases to contribute to the evolution of social norms and public policy, we must look for ways to connect deliberation among citizen-disputants to broader political processes and to encourage open (i.e., non-secret) settlements.

Broadly participatory negotiation processes have long been a major feature of the administrative rulemaking process. Consensus-building processes also are frequently used in efforts to resolve, or to contribute to the constructive management and eventual resolution of, the types of disputes that can give rise to significant cases but which are not necessarily the subject of pending litigation to which participants in the process are party. These processes, which often are facilitated by non-lawyer conflict resolution practitioners, have been used to foster understanding and facilitate cooperative action among opposing parties in a wide variety of disputes, including opposing factions in the abortion debate, policy makers and activists concerned with development and population policy, government officials, activists and taxpayers with opposing perspectives on the use of public funds for HIV/AIDS treatment programs, and environmentalists, landowners, government officials and others divided over how to respond to pollution and other public health problems.

Although, as Abram Chayes first observed, courts have become increasingly involved in helping parties fashion negotiated settlements, such broadly inclusive, consensus-based processes typically have not

of ordinary citizens, sometimes can be linked effectively to more inclusive political processes.


257. See, e.g., Chasin, supra note 219, at 327 (describing deliberative process involving opposing factions in the abortion debate); LeBaron & Carstarphen, supra note 139, at 1037–43 (describing deliberative process involving opposing factions in the abortion debate).


260. See, e.g., SUSSKIND & FIELD, supra note 103, at 42–59 (describing consensus building approach to resolving an environmental dispute).
been formally integrated into the adjudicative process. Over the course of the past decade or so, however, courts have increasingly begun to experiment with notions of problem-solving and restorative justice in ways that involve the disputants and other affected stakeholders in shaping the outcome of a judicial proceeding.

Margaret Farrell offers one vision of how litigation and negotiation could have interacted to produce a substantive outcome in Roe v. Wade that, in her view, may have more fully reflected the complexity of the issues involved and better accommodated the full range of affected parties—not just women and the unborn, but health care providers, fathers, local, state and federal agencies, and others. Farrell contends that by trying to resolve the social issues raised by abortion technology through litigation, we have transformed real-life, contextual, relational, complex facts about abortion into a two-sided contest between generalized maternal rights to privacy and theoretical state interests in potential human life, a process that teaches us little about the moral and social problems that we seek to resolve.

Farrell envisions the use of court-connected processes geared toward achieving collaborative solutions to the fundamental problems underlying a dispute. The specific process Farrell would have had the trial court in Roe v. Wade employ is an adaptation of the negotiated rulemaking process employed in the development of administrative regulations. This process typically involves an assessment period during which affected parties are identified, followed by notice to affected parties, a fact-finding stage, and a facilitated, consensus-based process that culminates in the development of a draft regulation that is

---

261. See Chayes, supra note 11, at 1298–1302.
264. See generally Farrell, supra note 11, at 330–53 (discussing Roe v. Wade, 410 U.S. 113 (1973)).
265. See id. at 274.
266. See id.
Farrell imagines a court-connected process that incorporates many of these elements. After *Roe* was initiated, she would have had the Court orchestrate a process in which other affected parties were identified and joined, discovery and other fact-finding efforts were undertaken, and settlement talks occurred (though not necessarily facilitated by the judge). If a consensus solution had emerged, it then would have been expressed in a multi-stakeholder settlement agreement or in a proposed bill or administrative rule. If a consensus solution did not emerge, the judge would have rendered a decision that was informed and constrained by the totality of facts presented, the competing perspectives of a broad range of affected parties, and, of course, existing substantive law.

I am not necessarily advocating widespread deployment of Farrell’s proposed approach in all its particulars. Needless to say, many practical and legal issues would have to be resolved before the exact process imagined by Farrell could be utilized. Farrell’s process nonetheless illustrates that it is possible to imagine a court-connected settlement process with the potential to inform official political processes and influence legislation and administrative policy. Procedures like Farrell’s, and others we could imagine, can create “feedback loops” among the various social and political actors that contribute to social change. Of course, there are constitutional and practical limits to what the parties can accomplish by agreement in some cases. In the *Roe* case, existing Texas law banned abortions. The parties could not have agreed simply to permit the plaintiff to terminate her pregnancy. Settlements resulting from deliberative processes must be designed to influence official lawmaking processes whenever current law prevents conduct the parties wish to permit.

Even when a settlement process is not formally or informally linked to legislative or administrative policy-making processes, however, settlement of a significant case can contribute to the fund of social norms that influence behavior (including others’ litigation and settlement decisions), the development of public policy, and subsequent judicial

---

268. See id. at 325–30.

269. Farrell anticipates and addresses many of these issues. See id. at 325–43 (discussing procedures for inclusion of nonparties and fact-finding).

270. Where a settlement does become linked to an official lawmaking process—for example, by producing and advancing a draft rule or bill—one would hope that the process by which regulators or legislators consider it is as deliberative as the process by which the parties produced it.

Settling Significant Cases

decisions. As Robert Ellickson has observed, “[i]n a well-functioning civilization . . . informal rules [that contribute to the maintenance of social order] . . . are among the most magnificent of cultural achievements.” Settlements resolving significant cases will not always have the force of law (except among the parties), but they may have important social influence nonetheless.

To maximize a settlement’s contribution to the fund of “informal rules” praised by Ellickson, its terms arguably should be made public. This is undoubtedly true, yet the mere fact of settlement may alter the social atmosphere surrounding the dispute. A settlement presumably signals to (approving and disapproving) others that some minimal economy of shared values developed among the parties. A settlement suggests that the parties concluded that they inhabit a universe of shared meaning after all, even if they occupy distant corners of it. The simple confirmation that settlement possibilities exist is arguably a significant contribution to the ongoing development of social norms concerning the central issues in any dispute involving deeply held values. As Robinson and his colleagues explain, dialogue processes “free partisans of some illusions that they hold not only about their ideological adversaries but about their own side as well (i.e., illusions of homogeneity, moral consensus, and extremism), which would in turn make it easier for them to express their own dissenting views.” If partisans who are not involved in a settlement see that agreement is possible, they may be more inclined toward conciliatory interaction with those ideological adversaries who they themselves encounter.

I doubt that many settlements of significant cases would remain entirely secret in jurisdictions where confidential settlements are permitted, since public disclosure of at least some of the principal settlement terms is likely to be a key interest of one or more of the parties and a necessary byproduct of implementation of the settlement. While damage payments and other monetary expenditures may be one feature of the parties’ agreement in some cases, most settlements of

272. ELICKSON, supra note 18, at 184.
273. See Luban, Erosion, supra note 9, at 2648–58 (arguing against confidential settlements). For my purposes, mass tort cases, where defendants frequently wish to keep secret the amount of damages paid in settlement of the plaintiffs’ claims, are not “significant cases.” Much of Luban’s concern about secrecy arises in response to settlement of these cases. Id.
274. Robinson, supra note 108, at 416. See also Thompson & Gonzalez, supra note 72, at 98 (arguing that disputants will be more persistent and creative in their efforts to reach agreement if there are precedents for resolving ideological conflict through negotiation).
significant cases about school prayer, abortion, pollution entitlements or
environmental resource allocation, affirmative action, and the like are
unlikely to be premised primarily on financial terms. Nonetheless, I
generally favor policies limiting or prohibiting confidential settlements
where they are not necessary to protect trade secrets, preserve
established privileges, or serve other narrowly focused and beneficial
social purposes.275

In this Part, I have argued that settlements, like adjudication, can have
democratic value and make positive contributions to the evolution of
social norms. David Luban, a thoughtful critic (though not an ardent
opponent)276 of settlement, acknowledges this possibility, but he believes
that adjudication’s contributions to democracy are bound to be
qualitatively superior to settlement’s contributions. Luban associates
adjudication with what he calls the “public-life conception” of politics,
which is characterized by forms of deliberation in which individuals
work at “building consensus around ideals rather than getting the right
answer, and in discovering worthy ends in addition to efficient
means.”277 Although Luban’s later writing on settlement acknowledges
the potential of consensual dispute resolution processes to contribute to
social change by occasionally producing what he calls “revisionary
justice,” he nonetheless associates settlement primarily with a “problem-
solving conception” of politics, which is about finding the cleverest
means to one’s desired end.278

Luban acknowledges that a “lucky and skillfully-conducted mediation
might actually end in both parties giving according to their abilities and
receiving according to their needs,” but he believes that such moments

275. For an overview of arguments for and against placing limits on confidential settlements, see
the following white paper published by New England Legal Foundation: CHRISTINE HUGHES,
CONFIDENTIAL SETTLEMENTS (2003) (on file with author). Confidentiality of court documents and
settlement terms must be distinguished from confidentiality of settlement negotiations, which
should remain private to the extent necessary to protect the parties or sustain the process. Parties to
deliberative processes sometimes participate at significant personal risk to their reputations (and
even their physical safety), and they should determine whether and when to publicize their
participation. See infra Part VI.B. Of course, sunshine laws may sometimes prevent parties from
conducting their discussions privately.

276. Luban’s earlier writings on settlement are somewhat more negative and skeptical than his
later writings. Compare Luban, Quality, supra note 35, at 407–11 (arguing that alternative dispute
resolution processes are unlikely to contribute to social change), with Luban, Erosion, supra note 9,
at 2634 (acknowledging limited potential of consensual dispute resolution processes to contribute to
social change).

277. Luban, Erosion, supra note 9, at 2634.

278. Id.
Settling Significant Cases

will be rare and that we therefore should not expect consensual dispute resolution processes to contribute significantly to major social change. Luban believes that consensual dispute resolution processes will seldom produce revisionary justice because more powerful parties will refuse to use them, or will exit from them, if they perceive that neutrals are attempting to empower weaker parties to make more than incremental gains through the process. According to Luban, this limits the extent to which dispute resolution processes can produce outcomes that depart significantly from existing legal entitlements.

It may be true that sweeping social change will seldom be achieved through deliberative dispute resolution processes, yet excessive attachment to adjudication may cause our collective consensus-building capabilities to atrophy, which arguably makes attempts at collective deliberation even more difficult to undertake and even less likely to succeed. I believe it is wrong to associate adjudication primarily with the service of the public interest and the potential for significant social progress and to associate settlement primarily with the service of private interests and incremental change within the current system of legal entitlements. While it is true that all parties, including those with greater economic and social power, enter mediation, adjudication, or any other dispute resolution process with their own self-interested objectives, those who participate openly in deliberative processes may find that their own and others’ perceptions of self-interest shift, producing greater alignment between competing conceptions of justice at play in the dialogue and a greater willingness to serve “public” objectives through whatever outcome is produced. At their best, consensus-based processes aimed at settlement of the types of disputes that give rise to significant cases are a form of political activity consistent with the public-life conception of politics that Luban favors.

Fiss, unlike Luban, is unwilling to grant that settlement can play any role in promoting social justice. He argues that settlement allows society

279. Luban, supra note 35, at 408–09.
280. Luban acknowledges that even incremental gains may be preferable to no gains at all, when that is what one can expect from litigation. “[W]hatever virtue ADR programs possess will lie in the fact they are marginally better than existing alternatives . . . and not in the expectation that they achieve utopian justice.” Id. at 387.
281. Id. at 409.
282. Luban certainly acknowledges that such shifts can occur. Luban, supra note 35, at 398. He simply doubts that many powerful parties will be willing to put themselves in harm’s way. Id. at 413.
to “mask[] its basic contradictions.” This is true only with respect to settlements that are strategically motivated. Processes that promote moral deliberation and perspective change actually bring social contradictions to the surface, where they can be constructively examined and addressed by the parties themselves. Compared to these processes, litigation allows us to mask our contradictions—chief among which is the desire to impose our own values on others, even though we know we live in a world that must be capable of peacefully reconciling our own values with others’ values. Litigation, as I have argued, temporarily shifts the broader social negotiation that must occur to a different location in the social sphere where, by institutional design, a “resolution” appears more readily achievable. As Rosenberg’s research arguably demonstrates, however, a U.S. Supreme Court decision seldom, if ever, single-handedly produces a durable resolution of deep social conflict. In the end, each deep division must be bridged, if at all, through a meta-dialog in which the Court is an influential, but not an omnipotent, participant.

It is, of course, impossible to say, from “the public’s perspective,” whether settlement of a particular ideologically based dispute is preferable to a judicial decision, because citizens’ views on any contested moral issue are, by definition, diverse. At best we can identify a range of coherent moral perspectives on a given issue that are held by significant numbers of people. If those citizens most immediately affected by a dispute experience some degree of genuine perspective change as a result of their participation in a process of moral dialogue, and if their changed perspectives produce somewhat more constructive and mutually beneficial relations, I would likely respect the outcome of their process, even if it did not give full expression to my own “chosen ideals.” If I did not, and if I had standing and cause to sue, I obviously would be free to seek a judicial decision consistent with my ideals. Indeed, one of the potential costs of settlement of any significant case is

283. Fiss, supra note 2, at 1086.
284. As Habermas states,

[D]eliberative politics remains part of a complex society, which, as a whole, resists the normative approach practiced in legal theory . . . . [T]he discourse-theoretic reading of democracy has a point of contact with a detached social-scientific approach that considers the political system neither apex nor center nor even the structural core of society, but just one action system among others. On the other hand, because it provides a safety mechanism for solving problems that threaten social integration, politics must be able to communicate through the medium of law with all the other legitimately ordered spheres of action . . . .

HABERMAS, supra note 82, at 302 (emphasis in original).
that the benefits realized as a result of the time, effort and other resources expended in achieving it will be nullified by a subsequent decision on the merits in a case brought by parties who are similarly situated.

VI. THE MORAL LEGITIMACY OF SETTLING SIGNIFICANT CASES

Having discussed the practical ways in which parties and the public may benefit through settlement of a significant case, or the type of dispute that may give rise to one, I wish to address the normative dimension of settlement more explicitly. Some argue that “peace” by whatever name—for example, “participatory politics,” “social capacity to deal with tough problems,” or “consensus”—should not come at the expense of “justice” or moral rightness. From this perspective, a compromise reached by parties to a value-laden dispute is indicative of weakness of character and moral relativism—which, in its strongest form, is the belief that everyone’s values are as valid as everyone else’s values, so it is wrong to judge others or to try to make them comply with one’s own values. Greater public participation and greater social capacity to manage strong disagreement may be important benefits of deliberative forms of democratic practice, but are the decisions they produce morally unsound? Is peace achieved at the expense of justice?

As indicated above, I do not consider courts to be uniquely capable of making moral judgments. Courts, however, sometimes do make absolutist judgments, in the sense that they validate one party’s moral perspective to the exclusion of other perspectives. In contrast, settlement typically involves acceptance of an outcome that balances the perspectives and objectives of all parties, though such outcomes can, and often do, favor one party’s perspectives and objectives over those of others. This distinction—between the inevitability of accommodation in settlement and the possibility of an uncompromising judicial decision—naturally raises questions about the moral integrity of settlement relative to a judgment reached through adjudication.

In this Part, I anticipate and counter claims that an agreement resolving a significant case necessarily lacks moral integrity. I also present a set of principles for designing consensual dispute resolution.

285. See Fiss, supra note 2, at 1075; see also Coleman & Silver, supra note 13, at 108 (arguing that “we sacrifice justice for the sake of efficiency and peace” when we settle).

286. See David Wong, Relativism, in A COMPANION TO ETHICS 442, 442 (Peter Singer ed., 1993).
processes geared toward resolving disputes involving deep value differences. Fidelity to these principles, I argue, increases the likelihood that the outcomes of such processes will be morally sound (as well as socially desirable and durable). We should not be satisfied with the increased public participation and greater social capacity to manage strong disagreement that may result from deliberative forms of democratic practice unless we also believe that such practices are capable of producing agreements that are morally sound.

A. Moral Relativism?

I should note preliminarily that I do not consider strategic settlement to be indicative of moral relativism for the simple reason that neither party has acknowledged the legitimacy of the other’s perspective. In the Piscataway case, for example, Ms. Taxman presumably believed that her moral perspective had been sufficiently vindicated through her trial and appellate court victories, and that further litigation therefore was not justified in light of other things she valued—for example, privacy and financial security, both of which might be jeopardized through continued litigation.287 The Piscataway school district determined that the best way to advance its values was not to attempt to advance them before the U.S. Supreme Court at that time. Neither party’s decision to settle was motivated by deference to the other party or the other party’s perspective.

The issue of moral relativism is more pressing in the context of settlements resulting from collective moral deliberation. Those who consider such settlements to be indicative of moral relativism are, in a sense, correct, but one must distinguish between moral relativism in its strong and weak forms. One who embraces moral relativism in its strongest form (as defined above) would have difficulty ever justifying litigation or any other form of coercive engagement with someone who holds different views. For an extreme relativist, negotiation would be the only acceptable means for resolving moral conflict—if indeed one sought to advance one’s substantive values in the face of others’ resistance at all.

In the weak form of moral relativism, one has genuine confidence in

287. Had Ms. Taxman settled before trial, rather than after achieving significant success in court, her decision still would not necessarily indicate moral relativism. While one party or the other is likely to be the clear victor in litigation, each party may reasonably conclude that the costs of proceeding to trial are too great in light of one’s tolerance for risk.
one’s moral convictions, but one’s convictions include what philosopher David Wong calls the justification principle. The justification principle holds that it generally “is wrong to impose one’s views on another person unless one can justify them to him or her . . . . [It is] an ethic that values interaction between members of different cultures through mutual consent.” One finds firm support for the justification principle in both deontological and social contractualist moral theory. From a Kantian perspective, the justification principle follows from the notion that people are to be regarded as ends in themselves, not as means to satisfaction of one’s own objectives or preferences. From the perspective of social contract theory, the justification principle affirms the moral value of consent in human relations.

The justification principle is an inherent feature of the theory of deliberative democracy. It gives rise to a prima facie duty to obtain consent before interfering with others’ actions, but this duty exists alongside those that flow from one’s other moral convictions. Whenever the justification principle conflicts with one’s other convictions, one must make a judgment about whether it should be overridden—that is, whether unilateral, coercive action in service of another conviction one holds is justified under the circumstances. In the context of disputes involving deeply held values, these circumstances may include those where the other party refuses to negotiate and the moral dangers of delay are great, or where the other party is not negotiating in good faith. Settlement is not itself a primary goal; when it is achieved, it is merely the byproduct of the parties’ efforts to serve their respective moral...

288. As Wong explains, confidence in one’s convictions need not depend upon an absolute belief that “one’s morality is the only true or the most justified one.” Wong, supra note 286, at 449.
290. Id.
291. Id. at 181; see also Wong, supra note 286, at 448 (associating weak form of moral relativism with social contract tradition).
292. See, e.g., Joshua Cohen, Deliberation and Democratic Legitimacy, in The Good Polity 17, 22 (Alan Hamlin & Philip Pettit eds., 1989) (noting that reason-giving and a commitment to resolving disagreements “through free deliberation among equals” are features of deliberative democracy); Gutmann & Thompson, supra note 61, at 55 (arguing that deliberative democracy can be embraced only by citizens motivated to find mutually acceptable terms for social cooperation). The search for mutually acceptable terms of association is a feature of the social contract tradition more generally, as expressed in the work of theorists such as Locke, Rousseau, Kant, and Rawls. See, e.g., Kuflik, supra note 35, at 56–57 (associating notion of “justice as mutual accommodation” with social contract tradition).
293. See Wong, supra note 289, at 186.
convictions in a manner that satisfies the justification principle. The justification principle gives rise to a presumption in favor of negotiation, however, at least for the purpose of sincerely exploring whether and where common ground may exist. The justification principle does not dilute or undermine one’s other moral convictions; rather, it engenders a certain humility with respect to them, and a certain confidence (as opposed to smugness) that allows one to engage in open and honest dialogue about them with other sincere people willing to engage one in good faith.

The strong form of moral relativism, which is associated with a lack of convictions tending toward nihilism, is too often conflated with moral relativism in its weak form. For example, Fiss implicitly brands proponents of settlement as advocates of the strong form of moral relativism by accusing them of pursuing peace at the expense of justice and by insisting that settlement in the service of one’s ideals is not possible. However, the substantive moral principles implicated in a dispute, and the justification principle alike, are legitimate values, all worthy of defense. One can be a relativist in the weak sense without being amoral, and one can pursue peace without sacrificing justice.

294. I would not translate this presumption into a requirement that parties attempt to resolve an ideologically based dispute as a condition of filing suit or proceeding to trial—a requirement of which Fiss also would disapprove. See Fiss, supra note 76, at 1670–71.

295. See CARL COHEN, DEMOCRACY 182 (1971) (“One’s knowledge of the truth and justice of his principles, or his firm belief in them, is counterbalanced within the [democratic] community by others who also know, or firmly believe, that other principles, in direct conflict with his own, are just and true.”); Benditt, supra note 15, at 35 (arguing that an “attitude of humility” creates potential for a compromise resolution of a conflict of principles).

296. See Wong, supra note 286, at 449 (noting that the strong form of moral relativism has given the weak form a bad name).

297. See Fiss, supra note 2, at 1075. Coleman and Silver argue that, from the public’s perspective, it would be desirable to litigate all disputes if litigation costs were zero, because litigated outcomes are just outcomes. See Coleman & Silver, supra note 13, at 108. Their claim has some appeal with respect to most ordinary cases, where social norms and legal norms are aligned and broadly accepted, but it is questionable when applied to significant cases, where social norms are hotly contested, for at least two reasons. First, losers of significant cases and their sympathizers are even less likely to view the unfavorable decision as just than losers of ordinary cases. Second, from the public’s perspective, the only potential advantage of litigating all significant cases is the possibility that a recently established or reaffirmed legal norm will dampen further conflict. As we have seen, however, the court’s ability to resolve social conflict on morally charged issues unilaterally is limited.

298. See Fiss, supra note 2, at 1086.

299. See JOHN RAWLS, A THEORY OF JUSTICE 126 (1971) (“The circumstances of justice may be described as the normal conditions under which human cooperation is both possible and necessary.”); Menkel-Meadow, Pursuing Peace, supra note 10, at 1767 (arguing that we should
Settling Significant Cases

Fiss and others are too quick to associate firmness of principle with morality, and compromise with expedience and moral weakness. Settlements reached through truly deliberative processes are not sell-outs. We do not succumb to moral relativism in its strong form merely by challenging ourselves to understand, and, when we glimpse something of their logic and legitimacy, accommodate, others’ values through agreements that also acknowledge our own. Settlements resulting from moral deliberation are evidence that the parties recognize the moral and social costs of effectively outlawing one perspective or the other, and have concluded that justice may find a fuller expression under the circumstances in some arrangement that attempts to confront the full complexity of the situation. We can compromise asserted claims and favored policies without sacrificing the moral principles on which they are based.

As I indicated earlier, all proponents of deliberative democracy consider reason-giving (and reason-probing) to be a central feature of deliberative processes, but they differ regarding the extent to which the structure and integrity of the process itself validates, from a moral perspective, the reasons on which an outcome is premised. At one end of the spectrum is Habermas’s discourse theory, which seems to hold that any outcome produced by a process that satisfies certain procedural norms is, by definition, morally justified, provided the participants

“seek to achieve ‘peaceful’ coexistence, mutual understanding and justice simultaneously”).

300. Luban sometimes seems to fall prey to such dichotomizing, as in a hypothetical he offers in one of his articles:

Suppose half the people in a nation believe—with arguments—“To each according to his need,” while the other half believes “To each according to his work.” (The latter too have arguments.) Each finds the other’s principle unacceptable; eventually they compromise on “To each according to his work, unless his work does not suffice to meet his most basic needs: then we keep him afloat with transfer payments.” The compromise principle is wrong, even morally wrong, if either side is right, for it violates distributive justice. It is, moreover, a principle believed by no one in the society.

David Luban, Bargaining and Compromise: Recent Work on Negotiation and Informal Justice, 14 Phil. & Pub. Aff. 397, 415 (1985) (emphasis in original). Luban assumes his negotiators reached their decision solely for purposes of expedience, by logrolling issues at play in the negotiation. He could have just as easily imagined the negotiators genuinely influencing each other, resulting in genuinely changed perspectives. There is nothing morally incoherent about the third principle Luban’s negotiators produce, as Luban suggests. Luban’s hypothetical loses its force if we imagine that the negotiators’ third principle is not the product of superficial horse-trading around entrenched interests, but the product of a truly deliberative process through which some degree of perspective change has occurred.

301. See Benditt, supra note 15, at 27.

302. See supra Part V.
regard each other as “free and equal” and the outcome is consistent with that premise. At the other end of the spectrum are Gutmann and Thompson, who believe that procedural norms can help promote, but cannot alone ensure, the moral integrity of the outcome of a deliberative process. In their view, the outcome of a deliberative process lacks moral justification if it fails to protect certain basic liberties and opportunities of those affected by it.

Like Gutmann and Thompson, I believe the outcome of a deliberative process cannot be morally justified solely by efforts to adhere to procedural norms. Proponents of alternatives to litigation have been criticized—fairly, I believe—as sometimes being too focused on procedures and not sufficiently focused on the substantive integrity of the outcomes produced by the procedures they advocate. The sort of ideal discursive conditions imagined by Habermas and others will seldom, if ever, be fully realized in practice. In the context of a process designed to explore the potential for settlement of a significant case, the participants and facilitator(s) will at best achieve a close approximation of the type of ideal process he imagines. If we accept that procedural norms cannot ensure the moral integrity of outcomes, however, a difficult question arises: Who determines whether the basic rights of those affected by a settlement are adequately protected?

In the United States, the decision about whether to settle, and on what terms, typically is wholly within the discretion of the parties themselves. Like Menkel-Meadow, I believe this is the proper place...
Settling Significant Cases

for decision-making power to reside. Fiss and other critics of settlement of significant cases would not, of course, deprive parties of the right to settle their dispute on terms of their own choosing. They simply do not recognize settlements as morally legitimate to the extent they reflect an apparent compromise of the parties’ moral convictions. This perspective is ironic, given that many judicial decisions often are careful “balancing acts” that try to accommodate features of competing moral views.

Of course, it is possible that parties to a process designed to encourage moral deliberation could agree to a settlement that they do not regard as even minimally protective of some basic liberty or opportunity, but which they regard as superior to the outcome they would achieve through litigation or legislation. Settlement is considered the least bad alternative, rather than a good alternative. In this event, while it is possible that parties could view settlement as pragmatic and rational in some very narrow sense, it is not morally justified according to the principles of deliberative democracy articulated by Gutmann and Thompson. Disengaging from the negotiation and pursuing one’s objectives through other means—including, perhaps, litigation, even if one fully expects to lose—is the morally preferable course of action, even if its effects are largely expressive. Settlement on terms that the weaker party does not consider adequately protective of basic liberties and opportunities achieves neither justice nor peace (in any sense that is likely to prove durable), so we obviously should not encourage it.

These arguments in favor of the justification principle and, implicitly, for moral relativism in its weak sense obviously will be entirely unsatisfying to anyone whose perspective on a charged moral issue is completely deontological—for example, a pro-life advocate who does not believe that a basic liberty of a woman who wishes to have an

---


307. See Fiss, supra note 2, at 1086 (“To settle for something means to accept less than some ideal.”).

308. See supra Part II.B.

309. See Gutmann, supra note 169, at 8 (“An immoral negotiated resolution to a moral conflict in politics may be worse than going to war or no resolution at all.”); see also Benditt, supra note 15, at 29–30 (arguing that genuine compromise cannot occur where “the alternative is to be forced into a still worse situation”). But see Carens, supra note 67, at 130 (“If it might be necessary for the party with the more legitimate interests to compromise at times simply because the other has more power, but it would be a political and moral error to accept the other party’s interests as legitimate simply because of this power.”).
abortion will be denied if she cannot obtain one, and that this interest should be afforded some degree of deference. Anyone who holds such views would react disapprovingly to a settlement in which, for example, a pro-life and a pro-choice organization agreed to an extended period of cessation from litigation on the condition that each organization engage in specific activities that they hope will measurably reduce both the number of unwanted pregnancies and the number of abortions. During this standstill period, abortions obviously would continue to occur, even if the parties’ efforts were dramatically successful at reducing their number. The only way a pro-life advocate could regard such an agreement as having moral integrity is through a willingness to accept a balancing of moral perspectives and objectives that he or she hopes will contribute to a reduction in unwanted pregnancies and increased utilization of alternatives to abortion (i.e., parenthood with adequate social and economic support or adoption), while also preserving abortion as an option that is used less frequently as a result of the parties’ efforts.\footnote{Such a balancing achieved through genuine moral deliberation differs from an outcome based upon utilitarian principles, even when the outcome resembles one that a utilitarian policymaking process might produce. As indicated above, deliberative processes seek to do more than uncover preferences and other private information in a search for an outcome that will maximally satisfy the participants’ competing interests. They seek to align preferences through discourse about the type of society in which the participants wish to live, and they do not necessarily treat all conceptions of the good as fungible, which much of utilitarian theory tends to do. See GUTMANN & THOMPSON, supra note 61, at 165–98 (comparing and contrasting deliberative and utilitarian approaches to policymaking).} The pro-life advocate who refuses to recognize the choice to terminate a pregnancy as a fundamental liberty must bear in mind, however, that the outcome of litigation may be a balancing of perspectives and objectives that he or she finds similarly unacceptable.

In sum, I find it difficult to understand why a decision reached by the disputants through an adequate deliberative process lacks moral integrity if a decision by empanelled judges does not—even if we assume, perhaps somewhat generously, that the interaction among appellate judges typically has the character of genuine moral deliberation.\footnote{For one recent example of interaction among appellate judges that falls far short of the deliberative ideal, see Adam Liptak, \textit{Order Lacking on a Court: U.S. Appellate Judges in Cincinnati Spar in Public}, N.Y. TIMES, Aug. 12, 2003, at A10 (reporting repeated accusations of “lying and underhanded conduct in important cases involving the death penalty and affirmative action” among the judges of the United States Court of Appeals for the Sixth Circuit).} For the reasons indicated at the beginning of this subpart, it also is difficult to understand why a strategic settlement like the one that concluded the \textit{Piscataway} litigation lacks moral integrity. Why should we not regard
Settling Significant Cases

an outcome that emerges from the types of negotiations I have described as having moral substance, particularly when adjudication, if it is presumed to have moral integrity, is an effective alternative to negotiation? As I have indicated, I accept the prerogative of parties to press their claims to a final judgment, with or without first attempting to negotiate, when they reasonably believe, after efforts to achieve consensus (or giving due consideration to such efforts), that there is only one way that justice can be served. I also believe, however, that the very fact that parties have this prerogative should dispel any general hesitation one might have about exploring the potential for a consensual resolution of the dispute through an adequate deliberative process.

B. Deliberative Democratic Theory and Dispute Process Design

If the structure of a deliberative process cannot ensure the moral integrity of the outcome it produces, sound design and administration of the process according to the principles of deliberative democratic theory can at least help promote the just, consensual resolution of disputes involving deep moral disagreement. Deliberative democratic theory provides meta-principles that dispute resolution practitioners can use to structure and manage dialogues intended to contribute to the resolution of disputes involving divisive moral issues. These principles are distinct from the many micro choices and moves that neutrals must make, but they can serve as touchstones that guide their process choices.

The foundational principles of deliberative democracy are concerned with its goals, its procedural features, and the ideal disposition of the participants in deliberative processes. These principles can be grouped and labeled, and their requirements elucidated, as indicated in the following chart. To the extent these principles do not encompass other, more familiar principles that are widely considered to be cornerstones of mediation and consensus-building processes—such as facilitator neutrality, informed consent, and party self-determination—they are intended to complement, rather than displace, them.

312. Habermas might object to the extension of deliberative principles into the realm of settlement processes, though Cohen, Gutmann, and Thompson would not. See GUTMANN & THOMPSON, supra note 61, at 131; HABERMAS, supra note 82, at 304–05; Cohen, supra note 292, at 21.
### PRINCIPLES FOR DELIBERATIVE DISPUTE RESOLUTION

#### PROCESS DESIGN

<table>
<thead>
<tr>
<th>Principle</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GOALS</strong></td>
<td></td>
</tr>
<tr>
<td>1. Reasoned Agreement</td>
<td>Consensus based on reasons that all parties recognize as being grounded in legitimate moral visions which are sincerely held and advocated in good faith.(^{313})</td>
</tr>
<tr>
<td>2. Protection of Basic Rights</td>
<td>All parties believe the agreement adequately protects basic liberties and opportunities.(^{314})</td>
</tr>
<tr>
<td>3. Public Influence</td>
<td>To the extent practicable, the process and/or its output are formally or informally linked to official political processes or otherwise designed to influence the development of social norms regarding the subject matter of the dispute.</td>
</tr>
<tr>
<td><strong>PROCEDURE</strong></td>
<td></td>
</tr>
<tr>
<td>4. Inclusiveness</td>
<td>The process is broadly inclusive of those potentially affected by its outcome or representatives who are accountable to them, all of whom</td>
</tr>
</tbody>
</table>

\(^{313}\) See Gutmann & Thompson, *supra* note 61, at 57; Habermas, *supra* note 82, at 306; Cohen, *supra* note 292, at 23.

\(^{314}\) See Gutmann & Thompson, *supra* note 61, at 199–29. Gutmann and Thompson do not address the difficult problem of who decides whether rights are minimally satisfied. In the context of settlement of significant cases, or the type of disputes that could give rise to them, I place that determination with the parties themselves.
Settling Significant Cases

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Publicity</td>
<td>The process and its outcome are made as public as they can be without compromising the physical or psychological security of the participants or the sustainability of the process itself.(^{316})</td>
</tr>
<tr>
<td>6. Open Agenda</td>
<td>Participants are free to make any assertion and to introduce any issue that could be addressed effectively through an agreement or could influence its terms.(^{317})</td>
</tr>
<tr>
<td>7. Open Exchange</td>
<td>Participants exchange and critically evaluate perspectives, information, proposals, and reasons.(^{318})</td>
</tr>
<tr>
<td>8. Reciprocity</td>
<td>Participants regard one another as free and equal persons and treat each other respectfully,(^{319}) keeping strategic motivations and coercive behavior in check, speaking to make oneself understood, and</td>
</tr>
</tbody>
</table>

---

315. See Gutmann & Thompson, supra note 61, at 128–64; Habermas, supra note 82, at 305; Habermas, supra note 303, at 89.
316. See Gutmann & Thompson, supra note 61, at 95–127; Habermas, supra note 82, at 305; see also Luban, supra note 35, at 416 (“The publicity principle—intended as a moral requirement on public policy—requires of any defensible policy that it be capable of withstanding general, public knowledge that it is in place.”).
317. See Habermas, supra note 82, at 306; Habermas, supra note 303, at 89.
318. See Habermas, supra note 82, at 305; Cohen, supra note 292, at 22.
319. See Gutmann & Thompson, supra note 61, at 52–94; Habermas, supra note 82, at 305–06; Cohen, supra note 292, at 23.
Washington Law Review  
Vol. 79:881, 2004

<table>
<thead>
<tr>
<th>Principle</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. Reflection</td>
<td>Participants are willing and able to express and jointly reflect upon the needs, attitudes and assumptions that underlie their own and others’ ideals, interests, and preferences.321</td>
</tr>
<tr>
<td>10. Openness to Influence</td>
<td>Participants attempt to suspend attachments to pre-existing positions and established norms at least enough to remain open to others’ perspectives and proposals.322</td>
</tr>
</tbody>
</table>

Needless to say, complete fidelity to these deliberative principles will be difficult or impossible to achieve, both for neutrals and for those well-intended participants who strive to maintain the proper disposition throughout the settlement process.323 The closer the ideal they represent is approximated, however, the more integrity the outcome of a deliberative settlement process is likely to have from a moral perspective.324 The principles can operate as regulative ideals that practitioners and parties attempt to approach without expecting to fully realize them. Indeed, some dispute resolution theorists and practitioners are consciously employing some or all of these principles in the design, administration, and evaluation of dispute resolution processes.325 Of

---

320. Minimizing strategic behavior is inherent in the notion of mutual respect and creation of a forum where participants are free from coercion. I thank Herbert Kelman for the expression “speaking to be understood and listening to understand” (personal communication with author).

321. See HABERMAS, supra note 82, at 306; Cohen, supra note 292, at 23.

322. See Cohen, supra note 292, at 22.

323. If those engaged in deliberation employ a facilitator, it also is needless to say that the facilitator’s training, skill, experience, rational and affective capabilities, biases, and general level of maturity will affect the process for good or ill, influencing the extent to which the principles are realized.

324. See GUTMANN & THOMPSON, supra note 61, at 17 (“Deliberative democracy does not assume that the results of all actual deliberations are just. In fact, most of the time democracies fall far short of meeting the conditions that deliberative democracy prescribes. But we can say that the more nearly the conditions are satisfied, the more nearly justifiable the results are likely to be.”).

325. See, e.g., JOHN FORESTER, THE DELIBERATIVE PRACTITIONER: ENCOURAGING
course, some dispute resolution practitioners undoubtedly utilize process designs and norms that satisfy some or all of the principles of deliberation without conscious appropriation of deliberative democratic theory.

Some principles of deliberation will be harder to remain faithful to than others in the context of settlement discussions among parties to a significant case. The principle of inclusiveness will be difficult to honor in many two-party disputes—particularly those among private parties, but even in those where the public is nominally represented through a government official or agency—unless the parties are willing to include representatives of other stakeholders, and other members of their social groups who hold different views, in the process (or unless litigation is pending and others have standing to intervene). The principle of publicity potentially is compromised if parties want the process to remain private, though the principle should acknowledge exceptions where secrecy is necessary to protect the security of the participants or the sustainability of the process. The less the outcome of the process will affect non-participants, however, the less concerned we should be about lack of inclusiveness and the secrecy of the process (as opposed to its outcome, which we should hope will be made public, so that it influences others’ perspectives and actions).

---

326. Gutmann and Thompson recognize that it sometimes may be appropriate for deliberation to occur secretly so that participants can “speak candidly, change their positions, and accept compromises without constantly worrying about what the public and the press might say.” GUTMANN & THOMPSON, supra note 61, at 115. Conflict resolution practitioners who facilitate dialogues among identity groups also recognize that secrecy may be necessary at some stages in the process to safeguard the reputations, and even the physical well-being, of those courageous enough to participate. See, e.g., Herbert C. Kelman, The Interactive Problem Solving Approach, in MANAGING GLOBAL CHAOS: SOURCES OF AND RESPONSES TO INTERNATIONAL CONFLICT 501, 507 (Chester A. Crocker et al. eds., 1996) (discussing value of secrecy in dialogues among representatives of groups engaged in conflict).
VII. CONCLUSION

It is ironic that the courtroom is viewed by many disputants, lawyers, and legal academics as the only acceptable forum for addressing disputes involving deeply held values, because litigation, like negotiation, entails compromise. Litigation involves compromise in at least two respects. First, litigation is a lottery in which the substantive values a party seeks to defend, and which it considers absolute, may be wholly or partially discredited by the court. Second, litigation merely shifts the burden of negotiation to others.

Negotiation should be viewed as a legitimate alternative to litigation for processing and resolving disputes involving deep moral disagreements. People accept the compromise inherent in adjudication of ideological disputes because they value not only the substantive moral perspectives they seek to defend, but also the pacific resolution of the dispute itself. Settlements achieved through deliberative dispute resolution processes may benefit both the parties and society in ways that litigation cannot. From the public’s perspective, deliberative dispute resolution processes can multiply opportunities for democratic participation and help citizens build social capacity to resolve tough problems. Even where parties are incapable of engaging in genuine moral deliberation, however, settlement for strategic reasons sometimes may be a sensible alternative for parties to a significant case, and one that should not invite scorn. Litigation and negotiation each have a legitimate role to play in our nation’s moral discourse and the evolution of social norms. Litigation and negotiation are complementary, mutually reinforcing social processes, even in the realm of disputes involving deep moral disagreement. Settlement has an important and constructive role to play in the pursuit of justice, the development of social norms, and the strengthening of social bonds. Interactions between law and social norms, and between the processes and institutions that produce them, are complex and unpredictable. We should not assume that one institution or process always is superior to others for righting wrongs or producing social change.
FIRST THINGS FIRST: FEDERAL COURTS SHOULD DETERMINE THE LEGAL STATUS OF A LLOYD'S OF LONDON SYNDICATE BEFORE DECIDING THE SYNDICATE’S CITIZENSHIP FOR DIVERSITY PURPOSES

John M. Brust

Abstract: Lloyd’s of London provides a marketplace where groups of underwriters form syndicates to insure risk. The United States Circuit Courts of Appeals have split on the question of how to determine whether a federal court has diversity jurisdiction over a controversy involving Lloyd’s syndicates. In a diversity action, each party must have diverse citizenship from all opposing parties. Circuit courts disagree about which diversity of citizenship test applies to suits involving Lloyd’s syndicates. The Second, Third, and Sixth Circuits have applied the real party in interest test. This test looks only to the citizenship of the parties that have a real interest in the litigation and ignores the citizenship of nominal or formal parties. In contrast, the Seventh Circuit has applied the unincorporated association test to assess the citizenship of a syndicate. This test requires a court to look to the citizenship of each of the underwriters in a syndicate. This Comment argues that the circuit courts have failed to consider whether a syndicate is a formal legal entity, which is a determination that will control the test that the court must apply. To answer this question, courts should first decide which law to apply to the case. Second, they should determine the syndicate’s legal status under that law. Courts should then apply the appropriate test to determine the citizenship of the underwriters. If the applicable law does not recognize a syndicate as a legal entity, then the real party in interest test applies. However, if the applicable law recognizes a syndicate as a legal entity, then the unincorporated association test applies.

Lloyd’s of London began in Edward Lloyd’s coffee house in the 1680s. Since that time, Lloyd’s has provided a market where buyers and sellers of insurance come together to negotiate insurance policies. Lloyd’s operates its market by placing underwriters into groups known as syndicates. Lloyd’s assigns an agent who is an expert in the insurance business to manage each syndicate on behalf of all its members.

Suits involving a Lloyd’s syndicate frequently give rise to the challenge that a federal court lacks diversity jurisdiction because

1. ANTONY BROWN, HAZARD UNLIMITED 16 (2d ed. 1978).
3. Id.
4. Id.
diversity of citizenship is not complete. Diversity jurisdiction requires that all parties on one side of a dispute be of different citizenship from all parties on the other side of the dispute. Because syndicates usually consist of numerous underwriters with different citizenships, it is common to have parties of non-diverse citizenship on opposite sides of the litigation.

The United States Circuit Courts of Appeals disagree about which test to apply to determine the citizenship of the underwriters on a Lloyd’s of London insurance policy. Under one approach, several circuit courts consider the underwriters of a syndicate to be various parties before the court and apply the “real party in interest” test to determine citizenship. A court applying the real party in interest test ignores the citizenship of a group’s representative, and instead looks at the citizenship of the parties who have a real interest in the action. Courts that have applied the real party in interest test to Lloyd’s underwriters have reached different conclusions as to which parties are the real parties in interest. Under a second approach, one circuit court has applied the “unincorporated association test,” provided by the U.S. Supreme Court in Carden v. Arkoma Associates. Under this test, courts attribute to the syndicate the
Diversity Citizenship of Lloyd’s Syndicates

citizenship of each of its member underwriters.\textsuperscript{14}

However, before a court can decide which test to apply to Lloyd’s underwriters, it must determine whether a syndicate is a legal entity under the applicable law.\textsuperscript{15} If the applicable law does not provide a syndicate with legal entity status, then the underwriters appear before the court as various parties, and the court must apply the real party in interest test.\textsuperscript{16} On the other hand, if the applicable law does provide the syndicate with legal entity status to sue or be sued as an unincorporated association, then the unincorporated association test applies.\textsuperscript{17} In this case, each syndicate will possess the citizenship of all its members.\textsuperscript{18} Before a court can address the legal status question, it must decide which law controls the determination of the legal status of a Lloyd’s syndicate.\textsuperscript{19} However, none of the circuit courts that have considered the citizenship of a Lloyd’s syndicate have considered what law should determine the syndicate’s legal status.\textsuperscript{20} Although the circuit courts have applied English law,\textsuperscript{21} state law,\textsuperscript{22} or have not specified which law they are applying,\textsuperscript{23} the courts have not shown how they determine which law controls a syndicate’s legal status.\textsuperscript{24}

This Comment provides a new approach to assist courts in determining the citizenship of Lloyd’s syndicates. Part I describes the structure of Lloyd’s. Part II discusses available choice-of-law rules and provides an analysis of a syndicate’s legal status under English law and state law. Part II also describes the bifurcated approach taken by one

\begin{thebibliography}{99}
\bibitem{1} Carden, 494 U.S. at 187 n.1.
\bibitem{2} See infra Part II.
\bibitem{3} See infra Part IV.B.
\bibitem{4} R.H. Bouligny, Inc., 382 U.S. at 153.
\bibitem{5} Id.
\bibitem{6} Id.
\bibitem{7} Id.
\bibitem{8} See Carden, 494 U.S. at 187–88. Congress has provided an exception to the legal entity test for corporations. See 28 U.S.C. § 1332(c) (2000). A corporation is deemed to be a citizen of the state in which it is incorporated and the state in which it has its principal place of business. \textit{Id}. The U.S. Supreme Court expressly declined to extend this exception to other legal entities in the absence of the expressed intent of Congress. \textit{R.H. Bouligny, Inc.}, 382 U.S. at 153.
\end{thebibliography}

district court in determining the legal status of unincorporated associations. Part III details the complete diversity requirement and the two tests that circuit courts have used to determine whether complete diversity exists in suits involving Lloyd’s syndicates. Finally, Part IV argues that courts should adopt a three-step approach to determine the citizenship of Lloyd’s underwriters. Under this approach, courts should (1) determine which law controls a syndicate’s legal status; (2) determine the syndicate’s legal status under that law; and (3) apply the appropriate citizenship test based on the syndicate’s legal status.

I. LLOYD’S OF LONDON PROVIDES A MARKET WHERE UNDERWRITERS FORM SYNDICATES TO SELL INSURANCE

Contrary to popular belief, Lloyd’s of London is not an insurance company. Rather, Lloyd’s provides a marketplace for its members to sell insurance. Lloyd’s refers to its underwriting members as “names” and places stringent requirements on who can become a name. Names must prove that they have a minimum amount of net worth, have the sponsorship of two current names, take part in a formal interview process, and have approval by a Lloyd’s committee. In addition, each name must travel to England to execute a number of investment related contracts, which specifically provide that they shall be governed by the laws of England. Once approved for membership, names join one of the several available underwriting groups, known as “syndicates.” A syndicate is a group of one or more names who have come together to insure risk. Individual names do not have the authority to underwrite insurance policies on their own. Instead, they rely on the syndicate’s underwriting agent to subscribe to insurance policies on behalf of all names in the syndicate. The underwriting agent has the authority to

25. See Layne, 26 F.3d at 41.
26. Id.
28. Id. at 227–28.
30. Id. at 229.
31. Id.
32. See id. at 229–30.
Diversity Citizenship of Lloyd’s Syndicates

bind all names within the syndicate to the insurance policy and represent their interests in a judicial proceeding.33

The Lloyd’s market operates in a central room within the Lloyd’s building.34 In this room, underwriting agents represent their syndicates and wait for insurance brokers to stop by and pitch their insurance needs.35 Insurance brokers, who represent outside parties seeking to insure against various types of risk, move from syndicate to syndicate until they find a syndicate that is willing to underwrite some or all of their clients’ risk.36 The first underwriting agent to commit his or her syndicate to a portion of the insurance policy is referred to as the “lead underwriter.”37 The lead underwriter may have a particular area of expertise and may negotiate the terms of the policy before committing to insure a portion of the risk.38 If the lead underwriter has committed his or her syndicate to only a portion of the risk, then the broker will continue to pitch the policy to other underwriting agents until he or she has covered one hundred percent of the risk.39 Underwriting agents for other syndicates may then subscribe to any remaining percentage of the policy, as negotiated by the lead underwriter.40 The lead underwriter manages the insured’s claims and has the authority to represent the interests of all the syndicates that have subscribed to the policy in any judicial proceeding.41

Parties who participate in the Lloyd’s of London market must comply with a series of complex rules and regulations. When a party joins Lloyd’s of London as a name, he or she agrees to specific rules for the allocation of risk, liability, and profits.42 In addition, Lloyd’s of London strictly regulates the interaction between each party within the market and imposes fiduciary responsibilities on all parties involved by

33. See Certain Interested Underwriters at Lloyd’s v. Layne, 26 F.3d 39, 42 (6th Cir. 1994).
34. See BROWN, supra note 1, at 3–4.
35. Id. at 4.
37. Id. at 145.
38. Id.
39. Id.
40. Id.
42. Beauchesne, supra note 27, at 226–29.
requiring that all dealings be in the “utmost good faith.” Lloyd’s has also established regulations detailing appointment and dismissal of underwriting agents; fiduciary duties such as accounting, disclosure, and avoidance of conflict of interests; and the regulation of brokers.

Additionally, each name is severally liable for his or her proportionate share of the insurance policy. When becoming a member of Lloyd’s, names agree to be personally liable for their percentage of the risk. Thus, if a particular name is liable for one percent of the risk subscribed to by his or her syndicate, and the corresponding syndicate subscribes to fifty percent of the risk on an insurance policy, then that name would be severally liable for one half of one percent of the total risk insured by the policy.

II. THE LEGAL STATUS OF A SYNDICATE DEPENDS ON WHICH LAW APPLIES

The unique structure of Lloyd’s has led courts to disagree as to the legal status of a Lloyd’s syndicate. The U.S. Circuit Courts of Appeals have not addressed the question of which jurisdiction’s law controls a syndicate’s legal status. However, the federal district courts have provided some guidance in answering this question. A court must determine which jurisdiction’s law applies before it can determine the legal status of a group of people. Depending on which jurisdiction’s law applies, courts have reached various conclusions as to the legal status of a Lloyd’s syndicate.

43. Id. at 46.
44. Id. at 46–49.
45. Certain Interested Underwriters at Lloyd’s v. Layne, 26 F.3d 39, 42 (6th Cir. 1994).
47. See, e.g., Corfield v. Dallas Glen Hills L.P., 355 F.3d 853, 864 (5th Cir. 2003), cert. denied, 124 S.Ct. 2421 (2004) (treating a syndicate as a group of individuals with separate and distinct rights); E.R. Squibb & Sons, Inc., 160 F.3d at 937 (treating a syndicate as a group of individuals with separate and distinct rights); Ind. Gas Co. v. Home Ins. Co., 141 F.3d 314, 317 (7th Cir. 1998) (treating a Lloyd’s syndicate as an entity); Layne, 26 F.3d at 42–43 (treating a Lloyd’s syndicate as a group of various individuals who are represented by a common agent).
48. See infra Part IV.B.
49. See Roby v. Corp. of Lloyd’s, 796 F. Supp. 103, 105–06 (S.D.N.Y. 1992) (analyzing several choice-of-law rationales to determine which jurisdiction’s law controls a syndicate’s legal status).
50. Id.
51. Compare Bobe v. Lloyd’s, 27 F.2d 340, 345 (S.D.N.Y. 1927) (holding that syndicates are not unincorporated associations under New York law), aff’d per curiam, 27 F.2d 347 (2d Cir. 1928),
Diversity Citizenship of Lloyd’s Syndicates

A. Courts Have Applied Various Choice-of-Law Rules to Determine Which Law Controls the Legal Status of a Lloyd’s Syndicate

A court must determine what law controls the legal status of a group before it can apply that law to determine the group’s legal status. Courts have acknowledged at least four competing methods to determine which law controls the legal status of unincorporated groups that operate in multiple jurisdictions: contractual choice-of-law clauses, the law of creation, the significant relationship test or similar analysis, and Federal Rule of Civil Procedure 17(b). Although U.S. Circuit Courts of Appeals have not explicitly considered the choice-of-law question in the context of a Lloyd’s syndicate, at least one district court has analyzed which jurisdiction’s law controls the syndicate’s legal status.

Under the first method, courts defer to the choice-of-law clauses in contracts. These clauses allow parties to specify which state’s law controls their relationship, thereby avoiding the uncertain application of a foreign jurisdiction’s laws. However, several state legislatures have enacted statutes that specifically prevent use of choice-of-law clauses in certain insurance contracts. These statutes are generally consumer protection statutes intended to protect citizens from out-of-state insurance companies. Thus, a court will not honor choice-of-law


54. Id.

55. Id.

56. Id.


58. COUCH, supra note 57, § 24:3.

975
clauses in insurance contracts that fall within the purview of such a statute. 59

Under the second method, courts apply the common law rule that requires courts to look to the law of the state that created the foreign entity to determine its legal status to sue or be sued. 60 Courts have applied this rule to determine whether a foreign group is a legal entity. 61 They have also used it to determine whether a foreign business group comprises an unincorporated association. 62 One district court applied this rule to a Lloyd’s syndicate and determined that English law governs the question of a syndicate’s legal status. 63 The court reasoned that English law created the syndicate because English citizens comprise the majority of membership in most syndicates and because all membership documents are executed in England. 64

Under the third approach, courts engage in a choice-of-law analysis by applying the forum state’s choice-of-law rules. 65 When choosing which law controls the contractual relationship between parties, the current trend is for states to adopt the Restatement of Conflict of Laws’ “significant relationship” test. 66 However, some states continue to rely

59. Id.

60. See 7 C.J.S. Associations § 2 (1980); see, e.g., Carl Zeiss Stiftung v. VEB Carl Zeiss Jena, 433 F.2d 686, 698–99 (2d Cir. 1970) (looking to the laws of West Germany to classify a German private business foundation without stockholders as a legal entity); Sanchez v. Bowers, 70 F.2d 715, 717 (2d Cir. 1934) (looking to Cuban law to classify a “sociedad de ganancias,” a profit-seeking association between husband and wife, as a separate juristic person); Four Way Plant Farm, Inc. v. Nat’l Council on Comp. Ins., 894 F. Supp. 1538, 1545–46 (M.D. Ala. 1995) (applying Florida law to conclude that a Florida-based insurance association is an unincorporated association); Cal. Clippers, Inc. v. United States Soccer Football Ass’n, 314 F. Supp. 1057, 1068 (N.D. Cal. 1970) (applying New York law to conclude that a gaming committee organized and operated in New York is an unincorporated association); Perkins v. First Nat’l Bank of Cincinnati, 79 N.E.2d 159, 163 (Ohio Ct. Com. Pl. 1948) (looking to the law of the Philippine Islands to classify a “sociedad anonima,” a business association organized under Philippine law, as a legal entity).

61. See, e.g., Carl Zeiss Stiftung, 433 F.2d at 698–99 (looking to the laws of West Germany to classify a German private business foundation without stockholders as a legal entity); Sanchez, 70 F.2d at 717 (looking to Cuban law to classify a “sociedad de ganancias,” a profit-seeking association between husband and wife, as a separate juristic person); Perkins, 79 N.E.2d at 163 (looking to the law of the Philippine Islands to classify a “sociedad anonima,” a business association organized under Philippine law, as a legal entity).


64. Id.

65. Id. at 107.

66. See RESTATEMENT (SECOND) OF CONFLICT OF LAWS § 188 (1988); 16 AM. JUR. 2D Conflict
Diversity Citizenship of Lloyd’s Syndicates

on traditional choice of law rules such as the place of the contract’s formation,67 the place of performance,68 or the intention of the parties.69 Under the significant relationship test, courts consider a list of factors to determine which state has the most significant relationship to the contract.70 The test also provides that if the parties negotiate and perform the contract in the same state, then courts should apply that state’s law to the contract.71

Under the final method, courts look to Federal Rule of Civil Procedure 17(b) to determine which law controls a party’s capacity to sue or be sued in federal court.72 Rule 17(b) instructs courts to look to the law of the forum state to determine the capacity of an unincorporated association to sue in federal court under diversity jurisdiction.73 However, one district court distinguished capacity from legal status and held that Rule 17(b) does not confer legal status on an unincorporated group of persons.74

One district court has explicitly applied a choice-of-law analysis to determine the legal status of a Lloyd’s syndicate.75 In Roby v. Corp. of Lloyd’s,76 the U.S. District Court for the Southern District of New York concluded that English law governs the legal status of a Lloyd’s syndicate.77 In that case, several New York citizens who were names in Lloyd’s of London syndicates sued their respective syndicates for

---

67. Id. § 100.
68. Id. § 102.
69. See RESTATEMENT (SECOND) OF CONFLICT OF LAWS § 188(2) (1988). The factors include “(a) the place of contracting, (b) the place of negotiation of the contract, (c) the place of performance, (d) the location of the subject matter of the contract, and (e) the domicile, residence, nationality, place of incorporation and place of business of the parties.” Id.
70. Id. § 188(3).
71. 16 AM. JUR. 2D Conflict of Laws § 94 (2003).
72. FED. R. CIV. P. 17(b).
73. Id.
75. Id.
77. Id. at 105.
violations of various securities laws and the Racketeer Influenced and Corrupt Organizations Act. The syndicates moved to dismiss the actions on the grounds that syndicates are not legal entities and cannot be sued. The district court first analyzed whether English or New York law should apply to determine legal status. It concluded that English law should control this determination. The court applied three alternative rules to determine the controlling law: the choice-of-law clause in Lloyd’s investment contracts providing for the application of English law to resolve disputes between names, the general rule that the law that created the association shall control its legal status, and New York’s choice-of-law “interest analysis.” Applying English law, the court dismissed the claim because it concluded that syndicates are not legal entities under English law. As an alternative rationale, the court concluded that syndicates are not legal entities under New York law. In addition, the court distinguished an entity’s legal status from its capacity to sue or be sued. The court found Federal Rule of Civil Procedure 17(b) did not control because a party must have both legal status, under the applicable state law, and capacity, under Rule 17(b), to appear in federal court.

B. Courts Have Applied the Laws of Various Jurisdictions and Reached Conflicting Results as to the Legal Status of a Syndicate

Once a court determines which law controls a group’s legal status, it must look to the applicable law to determine whether a syndicate is an unincorporated association and whether an unincorporated association is a legal entity under the laws of that jurisdiction. Under English law,

78. Id. at 103–04.
79. Id. at 104.
80. Id. at 106–07.
81. Id. at 107.
82. Id. at 106.
83. Id. at 106–07.
84. Id. at 107. The court weighed such evidence as the contract’s execution in England and the fact that over eighty percent of the investors are citizens of England. Id.
85. Id. at 111.
86. Id.
87. Id. at 110.
88. Id. at 110–11.
89. See id. at 106.
Diversity Citizenship of Lloyd’s Syndicates

syndicates are unincorporated associations. However, English law does not provide Lloyd’s syndicates with a separate legal personality apart from the individual underwriters. United States courts have widely recognized that English law does not provide a Lloyd’s syndicate with the legal status to sue or be sued in its own name.

When applying the laws of different states, courts have reached differing conclusions about the legal status of a Lloyd’s syndicate. The general common law definition of an unincorporated association is “a body of persons acting together, without a charter, but upon the methods and forms used by corporations, for the prosecution of some common enterprise.” Some courts have applied state law to conclude that Lloyd’s syndicates are unincorporated associations, while other courts have applied a different state’s law and reached the opposite conclusion.

Even if a court classifies a Lloyd’s syndicate as an unincorporated


93. Compare Bobe v. Lloyd’s, 27 F.2d 340, 345 (S.D.N.Y. 1927) (holding that syndicates are not unincorporated associations under New York law), aff’d per curiam, 27 F.2d 347 (2d Cir. 1928), with Merchs.’ & Mfrs.’ Lloyd’s Ins. Exch. v. S. Trading Co., 205 S.W.2d 352, 354–55 (Tex. Civ. App. 1918) (stating that a Lloyd’s association of underwriters constitutes an unincorporated association under Texas law), rev’d on other grounds, 229 S.W.2d 312, 316 (Tex. Comm’n App. 1921) (adopting judgment), and Youell, 203 F.R.D. at 509 n.6 (finding that a Lloyd’s syndicate meets the common law definition of an unincorporated association).

94. 7 C.J.S. Associations § 2 (1980).

95. See, e.g., Youell, 203 F.R.D at 509 n.6 (applying Kansas law to conclude that syndicates are unincorporated associations).

96. See, e.g., Bobe, 27 F.2d at 345 (applying New York law to conclude that syndicates are not unincorporated associations).
association, it does not necessarily follow that the applicable state law will treat the syndicate as a legal entity.\footnote{Youell, 203 F.R.D. at 508–09 (finding that a Lloyd’s syndicate meets the common law definition of an unincorporated association but that Kansas does not recognize unincorporated associations as legal entities).} The traditional common law rule is that an unincorporated association is not a separate legal entity.\footnote{7 C.J.S. Associations § 2 (1980); 6 AM. JUR. 2D Associations and Clubs § 51 (1999).} The jurisdictions that follow this common law rule will deny an unincorporated association the right to represent the collective interests of its members in a judicial proceeding.\footnote{6 AM. JUR. 2D Associations and Clubs § 51 (1999).} In such a case, parties could either bring an action on behalf of all individual members, no matter how many there may be,\footnote{7 C.J.S. Associations § 41 (1980).} or bring an action in the name of a few members as representatives of all the members.\footnote{Id. § 43. The representative option is known as the doctrine of “virtual representation,” and it is generally available where the members of the association have “a common or general interest in the subject matter of the suit, or where the members are numerous and it is impracticable to bring them all before the court.” Id. The selected representative should be a member of the association because there is a presumption that a member will fairly represent the rights and interests of all members. Id.}

Many states have enacted statutes to opt out of the common law rule and provide unincorporated associations with legal entity status.\footnote{See ALA. CODE § 6-7-81 (1993); CAL. CIV. PROC. CODE § 369.5 (West 2004); OKLA. STAT. tit. 12, § 2017 (Supp. 2004).} These statutes will generally allow an unincorporated association to bring or defend a suit as an entity representing the collective interests of all its members.\footnote{7 C.J.S. Associations § 47 (1980).} These statutes apply to both foreign and domestic unincorporated associations, as foreign unincorporated associations doing business in the forum state may be sued under the forum state’s statute.\footnote{Id.; see, e.g., Four Way Plant Farm, Inc. v. Nat’l Council on Comp. Ins., 894 F. Supp. 1538, 1548 (M.D. Ala. 1995) (“To determine whether an unincorporated association is subject to suit, the court examines the law of the forum state.”).} However, statutes that grant unincorporated associations legal entity status are “purely local in their operation and are not binding in other states.”\footnote{7 C.J.S. Associations § 47 (1980).} Thus, the legal status of a foreign unincorporated association can vary depending on whether the legislature in the forum state has enacted such a statute.\footnote{See Four Way Plant Farm, Inc., 894 F. Supp. at 1548–49 (applying the statute of the forum state).}
Diversity Citizenship of Lloyd’s Syndicates

C. One District Court Adopted a Bifurcated Approach to Determine the Legal Status of a Foreign Unincorporated Association Under State Law

In *Four Way Plant Farm, Inc. v. National Council on Compensation Insurance*, the U.S. District Court for the Middle District of Alabama analyzed the legal status of a foreign unincorporated association through a bifurcated choice-of-law approach. Under this approach, the court denied diversity jurisdiction to a Florida unincorporated association. The court applied Florida law to classify an association of insurers as an unincorporated association and Alabama law to determine whether the association was a legal entity. In applying Florida law, the court reasoned that it must look to where the association was created to determine whether it was an unincorporated association. Florida has utilized the common law definition of an unincorporated association, and the court found that the association fit within that definition. Next, the court reasoned that it must look to the laws of the forum state to determine whether an unincorporated association is subject to suit under its laws. Under Alabama law, unincorporated associations have legal entity status, and the court therefore determined that the association of insurers was a legal entity.

Although the U.S. Courts of Appeals have not considered the question of how to determine the legal status of a Lloyd’s syndicate, the district courts have provided some insight. District courts have looked to various choice-of-law methodologies to determine which law controls the question of legal status. Depending on which jurisdiction’s law state, Alabama, to grant legal entity status to a Florida unincorporated association).

---

108. *Id.* at 1545–46, 1549.
109. *Id.*
110. *Id.* at 1545.
111. *Id.*
112. *Id.* at 1545–48.
113. *Id.* at 1548.
114. *Id.*
115. See infra Part IV.B.
117. See *Four Way Plant Farm, Inc.*, 894 F. Supp. at 1547–48; *Roby*, 796 F. Supp. at 105–06.
controls the question, courts have reached different conclusions as to a syndicate’s legal status. Courts have adopted either a single law analysis or have used a bifurcated analysis to classify a Lloyd’s syndicate and determine its legal status.

III. A SYNDICATE’S LEGAL STATUS DETERMINES WHICH DIVERSITY JURISDICTION TEST FEDERAL COURTS WILL USE IN A SUIT INVOLVING THE SYNDICATE

Diversity jurisdiction requires that all parties on one side of a dispute be of diverse citizenship from all parties on the other side of the dispute. Courts may have difficulty determining the citizenship of the legal persons that can appear as parties. To resolve this problem, the U.S. Supreme Court has developed two different tests: the real party in interest test and the unincorporated association test. The real party in interest test applies unless state law provides a group of persons with legal entity status, in which case the unincorporated association test applies.

A. Federal Subject Matter Jurisdiction Extends to Disputes Between Parties of Completely Diverse Citizenship

Congress has extended subject matter jurisdiction to allow federal courts to resolve disputes between “citizens of different States.” Diversity jurisdiction, as this provision is commonly known, has two requirements: the amount in controversy requirement and the diversity of citizenship requirement. The amount in controversy requirement places the burden on the plaintiff to plead an amount greater than $75,000.

120. See Four Way Plant Farm, Inc., 894 F. Supp. at 1548.
123. Id.
125. Id. The amount in controversy requirement places the burden on the plaintiff to plead an amount greater than $75,000. See Jack H. Friedenthal et al., Civil Procedure § 2.8 (3d ed.
Diversity Citizenship of Lloyd’s Syndicates

provides that a federal court may not exercise diversity jurisdiction over a case or controversy unless the plaintiff and the defendant are citizens of different states.\textsuperscript{126} When determining the citizenship of the parties, courts look to citizenship at the time the plaintiff filed the complaint.\textsuperscript{127} In cases where there are multiple plaintiffs or multiple defendants, a court applies the “complete diversity” rule.\textsuperscript{128} The complete diversity rule provides that each plaintiff must be a citizen of a different state than each defendant.\textsuperscript{129} Thus, diversity jurisdiction will not exist if one plaintiff is a citizen of the same state as one defendant.\textsuperscript{130}

\textbf{B. Diversity Questions Commonly Arise in Suits by or Against Lloyd’s Syndicates}

In lawsuits filed against Lloyd’s underwriters in federal court, the basis for federal jurisdiction over the case is often diversity of citizenship between the plaintiff and the names.\textsuperscript{131} A syndicate’s legal status determines the test that a court must apply to determine its citizenship for diversity purposes.\textsuperscript{132} On one hand, if a syndicate is not a legal entity, then courts apply the real party in interest test.\textsuperscript{133} To determine whether complete diversity exists under that test, courts examine the citizenship of only the real party in interest.\textsuperscript{134} Sometimes, courts have found that the real party in interest is the lead underwriter.\textsuperscript{135}

\textsuperscript{126} 28 U.S.C. § 1332.
\textsuperscript{127} Mollan v. Torrance, 22 U.S. (9 Wheat.) 537, 539 (1824).
\textsuperscript{128} Strawbridge v. Curtiss, 7 U.S. (3 Cranch.) 267, 267 (1806).
\textsuperscript{129} 13B CHARLES ALAN WRIGHT ET AL., FEDERAL PRACTICE AND PROCEDURE: JURISDICTION 2d § 3605 (2d ed. 1984).
\textsuperscript{130} Id.
\textsuperscript{133} Id.
\textsuperscript{134} Id.
\textsuperscript{135} See, e.g., Layne, 26 F.3d at 43 (determining that lead underwriters are the real parties in
Other times, courts have concluded that all the names are real parties in interest. On the other hand, if a syndicate is an unincorporated association with legal entity status, courts apply the unincorporated association test and look to the citizenship of all the names in a syndicate.

1. The Real Party in Interest Test Applies if a Syndicate Is Not Considered a Legal Entity

Federal courts apply the real party in interest test when “various parties” appear before them. Under that test, courts determine which parties are simply representing the other parties, and which parties have a real interest in the action. To be a real party in interest, a party must have a personal stake in the outcome the litigation. Under this test, courts ignore the citizenship of the representative parties when determining whether complete diversity exists. Instead, courts look to the citizenship of only the real parties in interest in the action.

The U.S. Courts of Appeals for the Second, Fifth, and Sixth Circuits have all applied the real party in interest test to determine whether they had diversity jurisdiction over suits involving underwriters on a Lloyd’s policy. In Certain Interested Underwriters at Lloyd’s v. Layne, the Sixth Circuit looked to Tennessee law for guidance to determine who the real party in interest was. Under Tennessee law, the court concluded the lead underwriter was the agent of the syndicate, and its subscribing

---

136. See, e.g., E.R. Squibb & Sons, Inc., 160 F.3d at 930–31 (determining that all names are real parties in interest).
137. Carden, 494 U.S. at 187 n.1.
138. Id.
139. Id.
140. 6A CHARLES ALAN WRIGHT ET AL., FEDERAL PRACTICE & PROCEDURE: CIVIL 2d § 1556 (1982).
142. Id.
144. 26 F.3d 39 (6th Cir. 1994).
145. Id. at 43.
Diversity Citizenship of Lloyd’s Syndicates

underwriters were undisclosed principals on the insurance contract.\(^{146}\) Because Tennessee law holds agents liable for contracts they enter into on behalf of undisclosed principals, the *Layne* court determined that the lead underwriter was the real party in interest.\(^{147}\) Likewise, the court held that the other underwriters in the syndicate were not real parties in interest because Tennessee law releases undisclosed principals from liability when only the agent is sued.\(^{148}\)

In *E.R. Squibb & Sons, Inc. v. Accident & Casualty Insurance Co.*,\(^{149}\) the Second Circuit also applied the real party in interest test, but did not look to any state’s law to determine the party in interest.\(^{150}\) The court held that all subscribing underwriters are real parties in interest when the lead underwriter is suing in a representative capacity.\(^{151}\) However, the court also analyzed whether the plaintiff could avoid having to account for the citizenship of all subscribing underwriters by suing the lead underwriter in his capacity as an individual underwriter, and not in his capacity as the underwriter representing the syndicate.\(^{152}\) The Second Circuit remanded the question of whether, under English law, the lead underwriter could be sued as an individual and not as a representative of the other underwriters.\(^{153}\) On remand, the district court concluded that the lead underwriter could properly be sued in his or her individual capacity.\(^{154}\)

The most recent circuit court to employ the real party in interest test is

\(^{146}\) *Id.* at 43–44.
\(^{147}\) *Id.*
\(^{148}\) *Id.*
\(^{149}\) 160 F.3d 925 (2d Cir. 1998).
\(^{150}\) *Id.* at 931.
\(^{151}\) *Id.*
\(^{152}\) *Id.* at 935–36. The court reasoned that Lloyd’s was neither an unincorporated association nor a “constructive entity,” although the court did not specify which jurisdiction’s law it applied to make that determination. *Id.* at 937. The court distinguished Lloyd’s from other unincorporated associations because unincorporated associations have contractual provisions that run horizontally from member to member. *Id.* In contrast, the court concluded that a Lloyd’s of London insurance policy is composed of vertical obligations that run only from the insured to each underwriter individually. *Id.*
\(^{153}\) *Id.* at 939. In addition, the Second Circuit asked the district court to consider whether the unjoined names would be indispensable parties to the action under Federal Rule of Civil Procedure 19(b). *Id.*
the Fifth Circuit in Corfield v. Dallas Glen Hills L.P.155 In Corfield, a syndicate filed a declaratory action in federal court against a Texas partnership seeking a determination of the parties’ rights regarding an insurance contract.156 The syndicate’s lead underwriter filed suit on his own behalf and on behalf of the syndicate’s other underwriters.157 After the partnership challenged the court’s diversity jurisdiction, the lead underwriter amended the complaint to sue only on his own behalf.158 Like the Squibb court, the Corfield court did not consult state law to determine the syndicate’s legal status.159 The court concluded that the only citizenship that matters is the lead underwriter’s when he sues in his capacity as an individual underwriter.160

2. The Unincorporated Association Test Applies if a Syndicate Is an Unincorporated Association

The unincorporated association test applies when one or more of the parties to the litigation is a group of individuals or businesses that has formed, or is deemed by state law to have formed, an “unincorporated association” with legal entity status.161 Unincorporated associations include joint stock companies,162 labor unions,163 and limited partnerships.164 To determine the citizenship of a limited partnership for diversity purposes, for example, courts consider the citizenship of each of the partnership’s general and limited partners.165

The U.S. Supreme Court’s decision in Carden v. Arkoma Associates166 illustrates the role that state law plays in determining

155. 355 F.3d 853 (5th Cir. 2003), cert. denied, 124 S.Ct. 2421 (2004).
156. Id. at 855.
157. Id.
158. Id. at 856.
159. Id. at 863.
160. Id.
162. Chapman v. Barney, 129 U.S. 677, 682 (1889). A joint stock company, which is now a purely historical entity, was a hybrid resembling both a corporation and a general partnership. Carden, 494 U.S. at 202 (O’Connor, J., dissenting).
165. Id.
Diversity Citizenship of Lloyd’s Syndicates

whether the real party in interest test or the unincorporated association test applies to a group of persons. The Court addressed the question of how to determine the citizenship of a limited partnership for diversity purposes. The majority and dissent disagreed about how many parties were before the Court and which citizenship test applied. Whereas the dissent argued that there were many parties, the majority concluded that only one party appeared before the court—the limited partnership. The majority reasoned that the limited partnership did not consist of individual real parties in interest because state law recognized it as a “single artificial entity.” The Court stated that it would apply the real party in interest test when “various parties [are] before the Court” and would apply the rule for unincorporated associations when a “single artificial entity” was before the Court. The Court concluded that the dissent incorrectly applied the real party in interest test because the dissent erred in determining the legal status of the parties.

In contrast to the Second, Fifth, and Sixth Circuits, which have applied the real party in interest test, the U.S. Court of Appeals for the Seventh Circuit has applied the unincorporated association test to determine the citizenship of a Lloyd’s syndicate. In Indiana Gas Co. v. Home Insurance Co., the court concluded that a syndicate should be treated as an unincorporated association and that the syndicate assumes the citizenship of each of its subscribing underwriters. The court rejected the Sixth Circuit’s reasoning and conclusion in Layne. Instead, it relied on the U.S. Supreme Court’s decision in Carden to articulate the general rule that “every association of a common-law jurisdiction other than a corporation is to be treated like a

167. See id. at 187 n.1.
168. Id. at 186.
169. Id. at 187 n.1.
170. Id.
171. Id. at 187–88.
172. Id.
173. See id.
175. 141 F.3d 314 (7th Cir. 1998).
176. Id. at 317–19.
177. Id. at 319.
In reaching this conclusion, the court described the structure of a Lloyd’s syndicate as a cross between a general partnership and a limited partnership. The court stated that the syndicate’s liability structure resembles that of a general partnership because every name faces unlimited liability for its share of the risk, but the syndicate’s management structure resembles that of a limited partnership because the lead underwriter is the active manager, while the subscribing underwriters have no management powers.

Thus, the U.S. Circuit Courts of Appeals disagree as to which test applies to a Lloyd’s syndicate. The appropriate test for citizenship often determines a court’s ability to assert diversity jurisdiction over the case because of the complete diversity requirement. Some courts have applied the real party in interest test to determine the citizenship of the Lloyd’s parties, but these courts have reached different conclusions as to which parties are the real parties in interest. Other courts have applied the unincorporated association test and have concluded that a syndicate inherits the citizenship of all its members. The U.S. Supreme Court should resolve the disagreement among the courts by announcing a definitive methodology for determining the citizenship of a Lloyd’s syndicate.

---

178. Id. at 317.
179. Id. at 316.
180. See id.
181. See Corfield v. Dallas Glen Hills L.P., 355 F.3d 853, 863 (5th Cir. 2003) (applying the real party in interest test), cert. denied, 124 S.Ct. 2421 (2004); E.R. Squibb & Sons, Inc. v. Accident & Cas. Ins. Co., 160 F.3d 925, 930–31 (2d Cir. 1998) (applying the real party in interest test); Ind. Gas Co., 141 F.3d at 317 (applying the unincorporated association test); Certain Interested Underwriters at Lloyd’s v. Layne, 26 F.3d 39, 42 (6th Cir. 1994) (applying the real party in interest test).
182. See supra Part III.A.
183. See Corfield, 355 F.3d at 863; E.R. Squibb & Sons, Inc., 160 F.3d at 930–31; Layne, 26 F.3d at 42.
184. Compare E.R. Squibb & Sons, Inc., 160 F.3d at 930–31 (holding that all underwriters are real parties in interest when the lead underwriter is a party to the action as a representative of all names), with Layne, 26 F.3d at 43 (holding that under Tennessee law only the lead underwriter is the real party in interest).
Diversity Citizenship of Lloyd’s Syndicates

IV. FEDERAL COURTS SHOULD DETERMINE THE LEGAL STATUS OF A SYNDICATE UNDER STATE LAW, THEN DETERMINE ITS CITIZENSHIP FOR DIVERSITY PURPOSES

Of the U.S. Courts of Appeals that have analyzed the citizenship of a Lloyd’s of London syndicate, none has first decided which law should determine whether syndicates are legal entities. Instead of beginning with a choice-of-law analysis or announcing a definitive rule, the courts have simply looked to English or state law without explaining why one or the other should apply. Other courts have simply declared whether syndicates are legal entities, without relying on English or state law. Because the U.S. Supreme Court’s decision in Carden implies that courts look to state or English law to determine a syndicate’s legal status, courts should adopt a three-step process. First, courts must determine which state’s law controls the syndicate relationship. Second, courts must apply that law to determine if a Lloyd’s syndicate is a legal entity. Finally, courts must apply the appropriate test to determine the citizenship of the parties for diversity purposes.

A. Carden Suggests a Three-Step Approach for Determining a Syndicate’s Citizenship

The three-part test evolves from the U.S. Supreme Court’s language in Carden. The Carden Court implicitly recognized a state’s authority to create artificial entities, and the Court appeared willing to apply the legal entity test to any such state-created entity. When a group of people with similar interests are before a federal court, the court must determine which jurisdiction’s law it should apply when deciding whether the group is a legal entity or various separate persons. The

186. See Corfield, 355 F.3d at 863; E.R. Squibb & Sons, Inc., 160 F.3d at 930–31; Ind. Gas Co., 141 F.3d at 317; Layne, 26 F.3d at 42.
187. Compare E.R. Squibb & Sons, Inc., 160 F.3d at 936 (applying English law), with Layne, 26 F.3d at 43 (applying Tennessee law).
188. See, e.g., Ind. Gas Co., 141 F.3d at 316–17 (concluding that a syndicate is a legal entity without specifying what law the court applied to reach that conclusion).
190. See id.
191. See id at 187.
Carden Court stated that the real party in interest test applies when “various parties [are] before the Court” and that the unincorporated association test applies when the parties comprise a “single artificial entity.” To provide a uniform analysis, courts should adopt a three-part test to apply to a Lloyd’s syndicate. Courts should first determine which state’s law controls the syndicate relationship; second, apply that law to determine if a Lloyd’s syndicate is a legal entity; and third, apply the appropriate test to determine the citizenship of the parties for diversity purposes.

1. **Step One: A Court Must Determine Which Law Controls the Legal Status of a Lloyd’s Syndicate**

First, a court must decide what law should control the determination of a syndicate’s legal classification and status. Given the disparate treatment under the alternative laws, the answer to this question will often determine the appropriate test for citizenship. A court could choose to honor the Lloyd’s agreements that members must sign upon joining Lloyd’s, which contain a choice-of-law clause that provides that English law shall govern disputes between members. However, a court could reason that these contracts only regulate the relationship among members of Lloyd’s and not the relationship between the insured party and the syndicate. Many states have enacted statutes that would trump choice-of-law clauses in the insurance contracts, and a court could interpret such statutes to provide that the local law shall govern the entire insurance contractual relationship, including the legal classification of an

---

195. Roby v. Corp. of Lloyd’s, 796 F. Supp. 103, 106 (S.D.N.Y. 1992) (“Upon acceptance into the Society of Lloyd’s, the individual executes a number of documents, among which are the premium trust deed, a member’s agent agreement, and a managing agent agreement. Each of the documents provides that it is to be governed by the law of England.”).
196. See id.
197. COUCH, supra note 57, at § 24:3.
Diversity Citizenship of Lloyd’s Syndicates

unincorporated group of insurers.  

Alternatively, a court could adopt the common law rule that the applicable law is determined by looking to the jurisdiction in which the association was created. This was the primary rationale behind the Roby court’s holding that English law applied, and the rationale of the Four Way Plant Farm court when determining the legal classification of the group of insurers. A court applying this test to a Lloyd’s syndicate would likely conclude that syndicates are created and continue to operate under English law because the Lloyd’s of London market is physically located in England, the majority of names are English citizens, and each name must travel to England to execute the investment contract. However, courts may have difficulty determining which jurisdiction “created” the unincorporated association when members reside and conduct business in different locations.

As a third alternative, a court could apply the choice-of-law rules of the forum state to determine which state’s law should govern the syndicate’s legal status. The Roby court applied a choice-of-law analysis as an alternative holding and found that the New York choice-of-law “interests analysis” directed the court to apply English law. However, this outcome is dependent on the forum state’s choice-of-law rules. In addition, Roby is distinguishable from cases involving a determination of legal status for diversity purposes because Roby involved a dispute regarding the investment relationship between the names and Lloyd’s of London. In contrast, diversity cases result from the operations of a Lloyd’s syndicate in the United States.

Finally, a court could look to Federal Rule of Civil Procedure 17(b) to

198. Even though a court may determine that local law applies, the court could apply the whole law of the jurisdiction in which it sits to conclude that English or some other jurisdiction’s law controls the legal status of a syndicate.
200. See id.
204. See Roby, 796 F. Supp. at 106–07 (applying New York’s choice-of-law rules to conclude that English law governs “because Lloyd’s is based in London and approximately 80% of the Names are English citizens”).
205. See id. at 104.
conclude that the forum state’s law applies to determine a syndicate’s legal status. 206 However, Rule 17(b) explicitly states that it is a rule to determine the capacity of a party to appear in federal court. 207 As the Roby court discussed, capacity and legal status are two distinct concepts, and a party must have both to appear in federal court under diversity jurisdiction. 208 This rationale makes sense given that the language of Rule 17(b) directs federal courts to different law depending on the legal status of the party. 209 If a court were to look to Rule 17(b) to determine legal status, its analysis would be circular. 210

2. Step Two: A Court Must Apply the Applicable Law to Determine a Syndicate’s Legal Status

Once a court determines which state’s law applies, it must apply that law to classify the syndicate relationship and determine a syndicate’s legal status. 211 To do so, a court must first classify a Lloyd’s syndicate as either an unincorporated association or a mere grouping of individuals. 212 Then, a court must determine the syndicate’s legal status based on the applicable law. 213 A court could either adopt the bifurcated choice-of-law approach advocated by the Four Way Plant Farm court, or it could classify a syndicate and analyze its legal status under the same law. 214

States may choose to define an unincorporated association by statute, 215 but many of the statutes provide a definition similar to the common law definition. 216 Courts that have analyzed state law to classify

---

206. FED. R. CIV. P. 17(b).
207. Id.
208. Roby, 796 F. Supp. at 110.
209. FED R. CIV. P. 17(b).
210. Rule 17(b) directs courts to the law of the forum state to determine the capacity of an unincorporated association. Id. However, a court cannot determine whether a group comprises an unincorporated association until it looks to the law of some jurisdiction.
211. Roby, 796 F. Supp. at 110.
212. See supra Part II.B.
213. See supra Part II.B.
214. See supra Part II.B–C.
216. 7 C.J.S. Associations § 2 (1980). An unincorporated association is defined as “a body of persons acting together, without a charter, but upon the methods and forms used by corporations, for
Diversity Citizenship of Lloyd’s Syndicates

a Lloyd’s syndicate have reached different conclusions as to whether a syndicate qualifies as an unincorporated association.217 The majority of courts that have considered the question have concluded that a Lloyd’s syndicate is an unincorporated association.218 In addition, courts in the United Kingdom appear to recognize that a syndicate is an unincorporated association.219

After a court has classified a syndicate as either an unincorporated association or a group of individuals, it must then consider whether the syndicate is a legal entity.220 Courts apply either the same law they used to determine whether a syndicate is an unincorporated association or, under the bifurcated approach in *Four Way Plan Farms*, the law of the local jurisdiction.221 If a court determines that a Lloyd’s syndicate is not an unincorporated association, then the members are merely individuals and do not comprise a legal entity.222 However, if a court determines that a syndicate is an unincorporated association, it must then determine whether an unincorporated association is a legal entity under the applicable law.223 Although an unincorporated association is not a legal

---


218. See *Youell v. Grimes*, 203 F.R.D. 503, 509 n.6 (D. Kan. 2001) (“[T]he majority of courts [have held that] Lloyd’s Syndicates are unincorporated associations.”).


220. See, e.g., *Roby v. Corp. of Lloyd’s*, 796 F. Supp. 103, 107 (S.D.N.Y. 1992) (applying New York law to conclude that a syndicate is not an unincorporated association under and therefore not a legal entity); *Four Way Plant Farm, Inc.*, 894 F. Supp. at 1548 (applying the forum state’s statute to determine the legal entity status of a Florida unincorporated association).

221. *Four Way Plant Farm, Inc.*, 894 F. Supp. at 1548.


223. The court may choose the bifurcated approach and apply the forum state’s law to determine legal entity status, or it may choose to apply the same law it used to determine the classification as
entity under common law,\textsuperscript{224} many states have enacted statutes that confer legal entity status on unincorporated associations.\textsuperscript{225} Therefore, for a court to find that a syndicate has legal entity status, it must determine that a Lloyd’s syndicate is an unincorporated association and that unincorporated associations are legal entities under the applicable jurisdiction’s law. Thus, the outcome of step two of the three-part analysis will vary depending on the choice-of-law determination in step one.

3. Step Three: A Court Must Apply the Appropriate Citizenship Test Based on the Legal Status of a Lloyd’s Syndicate

Finally, a court must apply either the real party in interest test or the unincorporated association test depending on its determination of the syndicate’s legal status. If a court determines that a syndicate is not an unincorporated association or is an unincorporated association but not a legal entity, then the court must apply the real party in interest test because the names will be considered “various parties” before the court.\textsuperscript{226} In such an action, the names of the syndicate will usually appear before the court through their collective agent, who is usually the lead underwriter.\textsuperscript{227} A court applying the real party in interest test will determine the citizenship of the parties based on the citizenship of each name represented and will ignore their collective representative.\textsuperscript{228} Courts have reached different conclusions as to whether the lead underwriter is the only real party in interest.\textsuperscript{229} In addition, if the

\textsuperscript{224} See Four Way Plant Farm, Inc., 894 F. Supp. at 1548 (determining that courts should apply the laws of the forum state); 7 C.J.S. Associations § 47 (1980) (stating that a court has the option to apply the same law as it did to determine legal status).

\textsuperscript{225} Id. § 42.

\textsuperscript{226} Carden v. Arkoma Assocs., 494 U.S. 185, 187 n.1 (1990) (noting that the real party in interest test applies when various parties are before the court).

\textsuperscript{227} If a court concludes that a Lloyd’s syndicate is an unincorporated association without legal entity status, then such representation is referred to as “virtual representation” and would be permissible because the names have a common interest in the subject matter of the suit. See 7 C.J.S. Associations § 43 (1980).

\textsuperscript{228} N. Trust Co. v. Bunge Corp., 899 F.2d 591, 594 (7th Cir. 1990).

\textsuperscript{229} Compare Certain Interested Underwriters at Lloyd’s v. Layne, 26 F.3d 39, 43 (6th Cir. 1994) (holding that under Tennessee law only the lead underwriter is the real party in interest), with E.R. Squibb & Sons, Inc. v. Accident & Cas. Ins. Co., 160 F.3d 925, 930–31 (2d Cir. 1998) (holding that all underwriters are real parties in interest when the lead underwriter is a party to the action as a
Diversity Citizenship of Lloyd’s Syndicates

underwriters are merely “various parties” before the court with separate and distinct interests, a court may allow an action against only a portion of the names with diverse citizenship.230

If a court determines that a Lloyd’s syndicate is an unincorporated association with legal entity status, then it must apply the unincorporated association test to determine the citizenship of the syndicate.231 Under this test, the syndicate will be considered a citizen of every state in which a name resides,232 and diversity jurisdiction will depend upon the citizenship of all names as compared to the opposing parties. If any name is a citizen of the same state as an opposing party, diversity jurisdiction is destroyed for lack of complete diversity.233

B. Courts that Have Addressed a Syndicate’s Citizenship Have Failed To Determine Its Legal Status Under Applicable State Law

When considering the citizenship of a group of Lloyd’s underwriters, the courts of appeals have failed to address the syndicate’s legal status under the applicable jurisdiction’s law. In Layne, the Sixth Circuit assumed that the underwriting members were various parties before the court and therefore were subject to the real party in interest test.234 The court recognized that the underwriting group constitutes an unincorporated association,235 but neglected to provide any rationale for that conclusion. Similarly, in Indiana Gas Co., the Seventh Circuit concluded that a syndicate is an unincorporated association, but did not

---

230. See E.R. Squibb & Sons, Inc., 160 F.3d at 937 (2d Cir. 1998) (suggesting that diverse names may be used in their individual capacity, thereby preventing the destruction of diversity jurisdiction). Federal courts have disagreed about whether names are indispensable parties as provided in Federal Rule of Civil Procedure 19(b). Compare Allendale Mut. Ins. Co. v. Excess Ins. Co., 62 F. Supp. 2d 1116, 1124 (S.D.N.Y. 1999) (holding that each underwriter in Lloyd’s Syndicate was an indispensable party), with E.R. Squibb & Sons, Inc. v. Accident & Cas. Ins. Co., No. 82 CIV. 7327JSJSM, 1999 WL 350857, at *11 (S.D.N.Y. June 2, 1999) (holding that not all names were indispensable parties), aff’d, 241 F.3d 154, 162 (2d Cir. 2001).


233. Id.

234. Layne, 26 F.3d at 42.

235. Id. at 41–42.
specify what law it used to make that determination. 236 In addition, the Second and Fifth Circuits attempted to distinguish the Lloyd’s syndicate relationship from other unincorporated associations based on a purely factual analysis. 237 The Second and Fifth Circuits reasoned that the contractual provision requiring a name to honor all judgments rendered against any other name is a vertical obligation that runs between each name and the insured party. 238 However, the courts did not specify what jurisdiction’s law they used to define an unincorporated association, which they called a constructive entity. 239

C. The U.S. Supreme Court Should Have Affirmed the Fifth Circuit’s Decision in Corfield Based on the Three-Step Approach

The U.S. Supreme Court denied certiorari to Dallas Glen Hills, L.P. v. Corfield, 240 the most recent case involving the issue of diversity jurisdiction and Lloyd’s of London. 241 Rather than deny certiorari, the Court should have applied the three-step test to affirm the lower court’s ruling that the parties have complete diversity of citizenship. Applying the three-part test would likely have led the Court to the same conclusion, albeit with a different rationale, and provided guidance to the lower federal courts on how to address this complex problem. The analysis below applies the three-part test to the facts of Corfield.

Step One: Under the first step, a court should apply Texas choice-of-law rules to conclude that Texas law controls the legal status of a syndicate doing business in Texas. 242 The Texas legislature has enacted a statute that provides that Texas law shall govern all insurance contracts where the parties are doing business in Texas. 243 Therefore, a court should ignore contractual choice-of-law clauses when determining which law to apply. A court should interpret this statute to require that it apply

236. See Ind. Gas Co. v. Home Ins. Co., 141 F.3d 314, 317 (7th Cir. 1998) (classifying a syndicate implicitly as an unincorporated association because it is not a corporation).
239. Id.
241. Id.
242. See Corfield, 355 F.3d at 856.
Diversity Citizenship of Lloyd’s Syndicates

Texas choice-of-law rules to determine which jurisdiction’s law should control the legal status of the syndicate. In addition, a court should dismiss the common law rule of looking to the law of “creation” as too ambiguous because it is often difficult to tell which law “created” an unincorporated association when its members reside in many different jurisdictions and when it operates throughout the world. Finally, a court should clarify the distinction between legal status and capacity in order to avoid the circular inquiry that would result if the court looked to 17(b) to determine the applicable law of legal status. 244 Under Texas choice-of-law principles, a court should engage in an analysis under the Restatement’s most significant relationship test. 245 In applying the significant relationship test, a court should weigh the relevant factors 246 to conclude that Texas law controls because the syndicate was doing business in Texas. 247

Step Two: If a court applies Texas law, it must conclude that the Lloyd’s syndicate is an unincorporated association. 248 The Texas legislature opted out of the common law rule by enacting a statute that provides an unincorporated association doing business in Texas with legal entity status. 249 Alternatively, a court could reach the same outcome by following the bifurcated approach in Four Way Plant Farms because it would apply English law to classify the syndicate as an unincorporated association and Texas law to determine that the syndicate is a legal entity. 250

244. The circular inquiry results because 17(b) directs courts to the law of the forum state to determine the capacity of an unincorporated association, but a court cannot determine whether a group comprises an unincorporated association until it looks to the law of some jurisdiction.

245. See Sommers Drug Stores Co. Employee Profit Sharing Trust v. Corrigan, 883 F.2d 345, 353 (5th Cir. 1989). The Texas legislature has enacted a statute that directs courts to look to the law of the state of incorporation to resolve issues involving the internal affairs of a corporation. See Tex. BUS. CORP. ACT ANN. art. 8.02 (Vernon 2003). However, insurance entities are specifically excluded from the scope of the statute. Id. art. 2.01(B)(4)(d).


249. TEX. REV. CIV. STAT. ANN. art. 6133 (Vernon 1970).

250. Under the rationale in Four Way Plant Farms, 894 F. Supp. 1538, 1548 (M.D. Ala. 1995), a court could apply the law of “creation,” arguably English law, to classify the syndicate as an unincorporated association and the law in which the district court resided, Texas law, to determine
Step Three: Because the syndicate is a legal entity, a court should apply the unincorporated association test from *Carden*.\(^{251}\) In its application of the unincorporated association test, a court should conclude that the Lloyd’s syndicate in *Corfield* is a citizen of the United Kingdom because it consists of one underwriter, a British citizen. For diversity purposes, the partnership is a citizen of Texas, Delaware, and New York, while the syndicate, which consists of one underwriter, is a citizen of Britain.\(^{252}\) Thus, based on the application of the three-part test, the U.S. Supreme Court should have affirmed the Fifth Circuit’s holding that complete diversity exists between the parties.

V. CONCLUSION

The unique institution of Lloyd’s of London has caused the U.S. Courts of Appeals to disagree on the appropriate analysis to determine whether a court has diversity jurisdiction over a case involving a Lloyd’s syndicate. Some courts have held that a Lloyd’s syndicate is merely a collection of individuals and therefore the court must consider only the real party in interest when determining diversity jurisdiction. Other courts have held that a Lloyd’s syndicate is an unincorporated association and therefore inherits the citizenship of every name. However, the courts of appeals have not provided an analysis of which law they used to determine the legal status of Lloyd’s syndicates. This problem could be resolved if the U.S. Supreme Court adopts a three-step analysis to determine the legal status of a Lloyd’s syndicate. Although outcomes may differ depending on the conclusion of the three-step test, the test would provide courts with a uniform approach to assess diversity jurisdiction.

---

252. *Corfield*, 355 F.3d at 859.
WHEN PRISONERS ARE WEARY AND THEIR RELIGIOUS EXERCISE BURDENED, RLUIPA PROVIDES SOME REST FOR THEIR SOULS

Anne Y. Chiu

Abstract: The Religious Land Use and Institutionalized Persons Act of 2000 (RLUIPA) prohibits state and local governments from substantially burdening a prisoner’s exercise of religion unless the government can show that its action is the least restrictive means of furthering a compelling governmental interest. Prior to RLUIPA, courts subjected prisoners’ claims of violations of their right to exercise their religion to a “rational-relationship” standard. Because RLUIPA ("the Act") places a “strict scrutiny” standard on government actions burdening prisoners’ religious exercise, the Act is a legislative accommodation of religion. Under Lemon v. Kurtzman, legislative accommodations violate the Establishment Clause if their primary effect is to advance religion. When applying the Lemon test, the U.S. Supreme Court has considered three factors particularly relevant in determining whether a legislative accommodation advances religion: the severity of the burden on religion that the accommodation seeks to relieve, the burden the accommodation imposes on third parties, and the scope of the accommodation. This Comment argues that the U.S. Supreme Court should hold that RLUIPA is constitutional because it lifts substantial burdens on religious exercise, does not burden third parties, and has a broad enough scope in light of its secular legislative purpose.

In the prison system, the government closely regulates the right of prisoners to exercise their religion. In order to exercise this right, prisoners must usually request a “departure from otherwise applicable policies.” The government cannot completely deny a prisoner’s constitutional right to exercise religion, and it cannot impinge upon a prisoner’s right to religious exercise without a rational penological justification. Yet some prisons “restrict religious liberty in egregious and unnecessary ways” and impose substantial burdens on prisoners in practicing their religious faiths.

3. Id. at 1892.
The Religious Land Use and Institutionalized Persons Act of 2000 (RLUIPA) subjects prisoners’ claims of substantial government-imposed burdens on religious exercise to a “strict scrutiny” standard. RLUIPA was enacted to protect the religious liberty of institutionalized persons. Previously, burdens imposed on prisoners’ constitutional rights were subject to the “rational-relationship” test from Turner v. Safley. In Turner, the U.S. Supreme Court concluded that the rational-relationship test was sufficient to ensure prisoners’ rights under the Free Exercise Clause of the First Amendment. In contrast to the deferential Turner standard, RLUIPA prohibits government officials from imposing substantial burdens on the religious exercise of institutionalized persons unless the burdens are the least restrictive means of furthering a compelling government interest.

Because RLUIPA provides persons with greater protection for religious exercise while in prison, it is a legislative accommodation of religion. A “legislative accommodation” of religion lifts a neutral, generally applicable burden on religion imposed by the government.

8. Id. § 2000cc-1(a); Cutter v. Wilkinson, 349 F.3d 257, 264–65 (6th Cir. 2003) (noting RLUIPA requires courts to apply “strict scrutiny” to substantial burdens on religious exercise).
11. Id. at 89. The Free Exercise Clause states that “Congress shall make no law . . . prohibiting the free exercise . . . .” U.S. CONST. amend. I.
14. “Neutral” means that the statute on its face is not specifically directed at religious practice. See, e.g., Employment Div. v. Smith, 497 U.S. 872, 877–78 (1990) (distinguishing statutes that explicitly discriminate against or target religious practices as the object of the law from those that incidentally burden religion but are otherwise valid laws).
15. This Comment employs the term “legislative accommodation” narrowly to refer to the lifting or removal of neutral government-imposed burdens on religion, religious organizations, or individuals. See, e.g., Corp. of the Presiding Bishop v. Amos, 483 U.S. 327, 338 (1987) (exempting religious organizations from federal prohibition against religious discrimination in employment practices). Sometimes the courts label this an “exemption,” although this is not the same as the term of art “tax exemption.” See, e.g., Smith, 494 U.S. at 890 (citing a legislative accommodation as a “nondiscriminatory religious-practice exemption”). While “legislative accommodation” may also mean any legislative attempt to meet the religious needs of an organization or individual regardless of whether a state-imposed burden exists, see, e.g., Bd. of Educ. v. Grumet, 512 U.S. 687, 706 (1994) (legislative attempt to accommodate for the educational needs of the handicapped children in a Jewish sect), this broader definition is not used here. Therefore, although in certain cases the courts may use the term accommodation to have the broader meaning, this Comment focuses on the
Constitutionality of RLUIPA

Legislators enact some legislative accommodations in order to bring a government actor or program into compliance with the Free Exercise Clause, while other accommodations go beyond what the U.S. Constitution requires in order to meet the special needs of religion.

Since RLUIPA’s enactment, the federal circuits have split over whether the Act violates the Establishment Clause of the First Amendment to the U.S. Constitution. The circuit split turns on whether RLUIPA’s primary effect is to advance religion in violation of Lemon v. Kurtzman. U.S. Supreme Court decisions utilizing the Lemon test employ three factors to determine whether the primary effect of a legislative accommodation is to advance religion. These factors are: (1) the severity of the burden on religion the accommodation seeks to relieve; (2) the burden the accommodation imposes on third parties; and (3) the scope of the accommodation. The United States Courts of Appeals for the Fourth, Seventh, and Ninth Circuits have held that the primary effect of RLUIPA does not advance religion. In contrast, the United States Court of Appeals for the Sixth Circuit has held that RLUIPA’s primary effect is to advance religion.
This Comment argues that RLUIPA’s primary effect is not to advance religion. Part I discusses the historical context and legal landscape that prompted the enactment of RLUIPA and the protection RLUIPA provides for religious exercise in prisons. Part II discusses the U.S. Supreme Court doctrine on legislative accommodation in Establishment Clause challenges, with particular focus on the three relevant factors from this legislative accommodation precedent. Part III details the circuit split regarding RLUIPA’s constitutionality under the Establishment Clause. Finally, Part IV argues that RLUIPA is a permissible legislative accommodation because it lifts substantial burdens on prisoners’ religious exercise, does not impermissibly burden third parties, and has an appropriate scope in light of its secular legislative purpose.

I. CONGRESS RESPONDED TO THE U.S. SUPREME COURT’S DILUTION OF THE FREE EXERCISE CLAUSE BY ENACTING RLUIPA

Over the past forty years, the U.S. Supreme Court’s Free Exercise Clause jurisprudence has undergone a dramatic shift, which led Congress to enact RLUIPA in 2000. In 1963, the Court determined in Sherbert v. Verner that courts should apply a strict scrutiny standard to neutral government-imposed burdens on religious exercise. In 1990, the Court in Employment Division v. Smith retreated from this position and reinterpreted Sherbert and its other precedent to hold that the Free Exercise Clause did not require such an exacting standard. The Court held that the Free Exercise Clause provided little, if any, constitutional protection for religious exercise from neutral government-imposed burdens. In response to this ruling, Congress repeatedly attempted to restore the protection for religious exercise that the Court had recognized for its decision. Id. at 264. It also doubted whether RLUIPA resulted in excessive entanglement between church and state, but did not reach this issue because it concluded that RLUIPA had an impermissible effect. Id. at 267–68. This Comment focuses only on the conflict regarding the “primary effect” prong, and assumes that the purpose and entanglement prongs are not at issue.

26. See Erwin Chemerinsky, Constitutional Law § 12.3.2.2–12.3.2.4 (2d ed. 2002) (describing Establishment Clause jurisprudence since 1960).
28. Id. at 403.
30. Id. at 878–79.
31. Id.
Constitutionality of RLUIPA

in Sherbert.32 RLUIPA is Congress’s most recent attempt, applying the Sherbert strict scrutiny standard to the religious exercise of prisoners.33 Prior to RLUIPA, Turner’s deferential rational-relationship test governed the review of religious exercise violations in prisons.34 RLUIPA makes it more difficult for prisons to impose neutral, generally applicable burdens on religious exercise and therefore constitutes a legislative accommodation of religion.35

A. The Free Exercise Clause of the First Amendment Protects Religious Exercise

In 1963, the Sherbert Court first articulated a robust interpretation of the Free Exercise Clause that exemplified the high value the Court placed on prohibiting government interference with an individual’s religious exercise.36 The Court held that the Free Exercise Clause necessitated a strict scrutiny review of neutral, generally applicable laws that impose burdens on religious exercise.37 This standard required the government to prove a compelling government interest to justify the burden.38 The Court consistently upheld this interpretation in the decades following Sherbert.39

However, in 1990, the Smith Court broke from the Sherbert rule by reinterpreting the requirements of the Free Exercise Clause to provide little, if any, constitutional protection for religious exercise burdened by

33. Compare 42 U.S.C. § 2000cc-1(a) (2000) (“No government shall impose a substantial burden on the religious exercise of a person . . . unless the government demonstrates that imposition of the burden on that person is in furtherance of a compelling government interest . . . ”), with Sherbert, 374 U.S. at 403 (concluding that a “compelling state interest” may justify an incidental burden on religion).
35. See id. at 265.
36. Sherbert v. Verner, 374 U.S. 398, 403 (1963); see also Smith, 494 U.S. at 894–96 (O’Connor, J., concurring) (explaining U.S. Supreme Court jurisprudence that requires compelling governmental interests to justify encroachments on religious liberty); Wisconsin v. Yoder, 406 U.S. 205, 215 (1972) (noting that “only those interests of the highest order and those not otherwise served” can justify infringements on free exercise of religion).
37. See Sherbert, 374 U.S. at 406.
38. Id. at 403.
39. See Smith, 494 U.S. at 907 & n.1 (Blackmun, J., dissenting) (stating that the Court had been “consistent and exacting” in developing the Sherbert standard and citing the cases that had applied it); CHEMERINSKY, supra note 26, § 12.3.2.2.
neutral, generally applicable laws. Thus, today the Free Exercise Clause does not require the government to demonstrate a compelling government interest to justify those burdens. Furthermore, the Free Exercise Clause does not require legislatures to accommodate religion by exempting religious individuals from those burdens. Instead, the Smith Court affirmed the power and the discretion of legislatures to enact statutory protection beyond what the Free Exercise Clause required. That is, the Court indicated that although the Free Exercise Clause did not require it, legislatures could enact accommodations to protect religion beyond the Free Exercise Clause minimum without “establishing” religion and thereby running afoul of the Establishment Clause.

B. Congress’s Legislative Response: RFRA and RLUIPA

In response to Smith, Congress enacted several pieces of legislation, including RLUIPA, to restore the Sherbert protection for religious exercise. The Religious Freedom Restoration Act of 1993 (RFRA) explicitly adopted the Sherbert strict scrutiny standard to protect religious exercise in all spheres. However, in 1997, the Court in City of Boerne v. Flores held RFRA unconstitutional, as applied to the states, because Congress exceeded its powers under the Fourteenth Amendment in enacting RFRA. The Boerne Court did not rule on the issue of the

40. Smith, 494 U.S. at 882–85. The Court noted some possible exceptions to this rule. One includes hybrid rights where free exercise involves other constitutional protections like freedom of speech; another includes contexts where there has been “individualized governmental assessment” of the reasons for the relevant conduct. Id. at 881, 884.
41. See id. at 888–89.
42. Id. at 878–80.
43. Id. at 890.
44. See id. The Court in Locke v. Davey, 124 S. Ct. 1307 (2004), also affirmed that there is legislative discretion to protect religious exercise at a greater level than required by the Free Exercise Clause, but within the level permitted by the Establishment Clause. Id. at 1311.
47. Id. § 2000bb-1.
49. Id. at 536. Although the Court’s language in Boerne is ambiguous, RFRA appears to be valid as applied to federal laws and practices. See, e.g., O’Bryan v. Bureau of Prisons, 349 F.3d 399, 400–01 (7th Cir. 2003) (determining that “RFRA may be applied to the internal operations of the national government”); Guam v. Guerrero, 290 F.3d 1210, 1221 (9th Cir. 2002) (stating that RFRA
Constitutionality of RLUIPA

constitutionality of the strict scrutiny standard under the Establishment Clause.50

In response to Boerne,51 Congress enacted RLUIPA (“the Act”). The Act, which applies to regulations involving land use and institutionalized persons, requires the Sherbert strict scrutiny of government actions in religious exercise claims.52 Because Congress passed RLUIPA using its powers under the Spending and Commerce Clause powers,53 the Act applies only to state and local prisons and prison officials that receive federal funds, as well as to their regulations that substantially burden religious exercise and affect interstate commerce.54

C. RLUIPA Limits the Burdens that States Can Impose on Prisoners’ Exercise of Religion

The U.S. Supreme Court has determined that religious exercise rights in prisons are more circumscribed than religious exercise rights outside prison walls.55 In 1987, when the Sherbert strict scrutiny standard still applied to religious exercise claims outside prisons, the Court in Turner established that courts should apply the rational-relationship test when determining whether a state action impermissibly infringes on a prisoner’s fundamental rights.56 Under this test, prison regulations burdening prisoners’ constitutional rights are valid as long as they bear a rational-relationship to penological interests.57 The rational-relationship

“is constitutional as applied in the federal realm”); Henderson v. Kennedy, 265 F.3d 1072, 1073 (D.C. Cir. 2001) (holding that the portion of RFRA applicable to the federal government “survived the Supreme Court’s decision striking down the statute as applied to the States”); Kikumura v. Hurley, 242 F.3d 950, 958 (10th Cir. 2001) (concluding that Boerne “does not determine the constitutionality of RFRA as applied to the federal government”); Christians v. Crystal Evangelical Free Church (In re Young), 141 F.3d 854, 856 (8th Cir. 1998) (concluding that RFRA is constitutional when applied to federal law).

50. See Boerne, 521 U.S. at 512, 536.
53. Id. § 2000cc-1(b).
54. Id.
56. The Law of Prisons, supra note 2, at 1893.
test extends considerable deference to prison officials.\footnote{58}{The Law of Prisons, supra note 2, at 1893.}

In \textit{O’Lone v. Estate of Shabazz},\footnote{59}{482 U.S. 342 (1987).} the Court affirmed that this rational-relationship test applies to prisoners’ religious rights.\footnote{60}{Id. at 349.} Although the \textit{O’Lone} Court held that the Free Exercise Clause only requires courts to apply a rational-relationship standard when reviewing violations of the religious exercise rights of prisoners,\footnote{61}{Id.} RLUIPA raised the standard to strict scrutiny.\footnote{62}{42 U.S.C. § 2000cc-1(a) (2000).} Under RLUIPA, state and local governments may not substantially burden the religious exercise of institutionalized persons unless the burden is the least restrictive means of furthering a compelling government interest.\footnote{63}{Id.}

In order to avoid or relieve burdens on religious exercise, the government may choose to enact a “legislative accommodation,” a statute enacted by the legislature to lift a neutral, generally applicable burden on religion imposed by the government.\footnote{64}{This Comment employs the term “legislative accommodation” narrowly to refer to the lifting or removal of neutral government-imposed burdens on religion, religious organizations, or individuals.} RLUIPA requires that regulations imposing a substantial burden on prisoners’ religious exercise rights satisfy a strict-scrutiny standard.\footnote{65}{42 U.S.C. § 2000cc-1(a). RLUIPA has the same strict-scrutiny standard as RFRA. See 42 U.S.C. § 2000bb-1 (2000).} Without the Act, government-imposed burdens on these rights would be subject to the rational-relationship review.\footnote{66}{See Cutter v. Wilkinson, 349 F.3d 257, 263 (6th Cir. 2003).} By holding prisons officials to this higher standard, RLUIPA makes it more difficult for prison officials to impose the burden on religion and therefore constitutes a legislative accommodation of religion.\footnote{67}{Id. at 263–64.}

\section{II. LEGISLATIVE ACCOMMODATIONS AND THE ESTABLISHMENT CLAUSE}

The U.S. Supreme Court’s Establishment Clause jurisprudence permits legislative actions to accommodate religion beyond the level
Constitutionality of RLUIPA

required by the Free Exercise Clause. However, the Court has determined that some legislative accommodations advance religion and therefore violate the Establishment Clause. To determine whether a legislative accommodation violates the Establishment Clause, courts apply the three-prong test announced by the U.S. Supreme Court in *Lemon v. Kurtzman*. Under the *Lemon* test, a legislative accommodation does not violate the Establishment Clause if it passes a three-pronged analysis: (1) it has a secular legislative purpose; (2) its primary effect is neither to advance nor inhibit religion; and (3) it does not foster excessive government entanglement with religion. The second prong of the test is at issue in decisions construing RLUIPA. This prong concerns whether RLUIPA’s primary effect is to advance religion in violation of the Establishment Clause. When considering the constitutionality of legislative accommodations under the *Lemon* test, the U.S. Supreme Court has considered three factors: (1) the severity of the government-imposed burden on religion that the accommodation seeks to relieve; (2) the burden that the


70. 403 U.S. 602, 612–13 (1971). See, e.g., Madison v. Riter, 355 F.3d 310, 316 (4th Cir. 2003) (stating that the *Lemon* test is the appropriate method to determine whether RLUIPA is constitutional); Cutter, 349 F.3d at 262–63 (6th Cir. 2003) (electing to use the “traditional three-part *Lemon* analysis” to evaluate RLUIPA’s constitutionality); Charles v. Verhagen, 348 F.3d 601, 610 (7th Cir. 2003) (noting that the *Lemon* test is the appropriate method to determine whether Congress violated the Establishment Clause); Mayweathers v. Newland, 314 F.3d 1062, 1068 (9th Cir. 2002) (explaining that the *Lemon* test determines whether an accommodation statute is neutral).


72. *Madison*, 355 F.3d at 318; *Cutter*, 349 F.3d at 267; *Charles*, 348 F.3d at 611; *Mayweathers*, 314 F.3d at 1068.

73. See *Lemon*, 403 U.S. at 612.

74. Although the U.S. Supreme Court ruled on some legislative accommodations prior to *Lemon*, it paid close attention to the primary effect of advancing religion prior to setting down the formal test in *Lemon*. See, e.g., Gillette v. United States, 401 U.S. 437, 450 (1981) (noting that the Establishment Clause requires a secular purpose, evenhandedness in operation, and a neutral primary impact, and prohibits undue government involvement in religious affairs); Walz v. Tax Comm’n, 397 U.S. 664, 668, 674 (1970) (considering whether statute’s purpose was to advance or inhibit religion and whether there was excessive government entanglement with religion).

75. See Texas Monthly, Inc. v. Bullock, 489 U.S. 1, 18 n.8 (1989) (plurality opinion) (finding permissible “legislative exemptions . . . that were designed to relieve government intrusions that might significantly deter adherents of a particular faith from conduct protected by the Free Exercise
accommodation places on third parties, and (3) the scope of the accommodation.

A. Legislative Accommodations Must Lift a Significant Burden on the Exercise of Religion

Under the Establishment Clause, Congress may accommodate significant burdens or “serious encroachments on protected religious freedoms” that result from the implementation of neutral laws of general applicability, even when the burdens do not “prohibit” the free exercise of religion and violate the minimum protection required by the Free Exercise Clause. A statute or regulation imposes a significant burden on individuals when it forces them to choose between acting in a manner offensive to their beliefs or suffering penalties. For example, in *Gillette v. United States*, the U.S. Supreme Court upheld a federal
Constitutionality of RLUIPA

conscientious objector statute that exempts from military service anyone who opposes all war by reason of “religious training and belief.” 82 The Court noted that while the Free Exercise Clause did not require the exemption of conscientious objectors from compulsory military service, 83 opposition to war is an issue of conscience and “duty to a moral power higher than the State.” 84 The Court concluded that forcing people opposed to war on religious grounds to choose between “contravening imperatives of religion and conscience or suffering penalties” would significantly burden these religious individuals. 85 Thus, the Court found that it was permissible for Congress to attempt to accommodate “free exercise values,” to avoid “unnecessary clashes with the dictates of conscience.” 86

A neutral, generally applicable law imposes a significant burden when it may inhibit the missions of religious entities due to their fear of liability. 87 In Corp. of the Presiding Bishop v. Amos, 88 the U.S. Supreme Court upheld § 702 of the Civil Rights Act of 1964 (CRA) as constitutional under the Establishment Clause. 89 Title VII of the CRA prohibits discrimination in employment on the basis of religion, 90 but § 702 exempts religious organizations from this rule. 91 While the Court assumed that the Free Exercise Clause did not require the accommodation, 92 it reasoned that without the § 702 exemption, religious organizations’ fear of liability for employment discrimination under Title VII would inhibit the way the religious organizations carry out their missions. 93 The Court concluded that this fear was a

82. Gillette, 401 U.S. at 450–51.
83. Id. at 461 n.23.
84. Id. at 453.
85. See id. at 445, 453 (noting congressional “deep concern” for imposing a choice between suffering penalties for following one’s religious conscience and following a law that “clashes” with one’s religious conscience). Although the issue in Gillette centered on the distinction between those who opposed particular wars from those who opposed all war, both based on religious grounds, id. at 447, the Court’s rationale highlights the kind of burden on religious conscience that the Court would deem warrants a legislative accommodation. See id. at 445, 453.
86. Id. at 453.
89. Id. at 330.
90. Id. at 329.
91. Id.
92. Id. at 336.
93. Id. Without § 702, the organizations would still be permitted to use religious criteria to hire
“significant burden” that Congress could accommodate.94

Unlike the burdens of offending religious conscience and inhibiting religious activity, the U.S. Supreme Court determined the payment of a sales tax is not a burden severe enough to justify a legislative accommodation.95 In Texas Monthly, Inc. v. Bullock,96 a plurality of the Court concluded that a state sales tax on periodicals was not a sufficient burden to justify exempting exclusively religious periodicals from the tax.97 The plurality noted that the tax break was not a proper legislative accommodation because the Free Exercise Clause did not require it and it could not reasonably be seen as removing a state-imposed “deterrent to the free exercise of religion.”98 It found that the Texas tax exemption did not remove a “demonstrated and possibly grave imposition on religious activity,”99 and that tax payment would neither offend religious beliefs nor inhibit religious activity.100

B. Legislative Accommodations Must Not Substantially Burden Third Parties

A legislative accommodation may not substantially burden third parties to a point of “unyielding weighting” in favor of the interests of the ones accommodated.101 For instance, the U.S. Supreme Court in

94. Id. at 335–36.
97. Id. at 15, 18 n.8 (plurality opinion).
98. Id. at 15 (plurality opinion). While the plurality opinion is not binding precedent, it should be noted that the Texas Monthly decision preceded Smith, which dramatically changed Free Exercise Clause jurisprudence. See supra Part I.A. Therefore, the Texas Monthly plurality most likely applied a pre-Smith understanding of the Free Exercise Clause requirements.
99. Texas Monthly, 489 U.S. at 18 n.8 (plurality opinion). While the Texas Monthly plurality characterized the Amos accommodation as a requirement of the Free Exercise Clause, the Amos Court assumed that it was not. Amos, 483 U.S. at 336.
100. Texas Monthly, 489 U.S. at 18 (plurality opinion). Although the plurality entertained the possibility that some religious groups might contend that tax payment violates their religious tenets, an overriding governmental interest would likely justify the tax burden on religion. Id. at 19–20 (plurality opinion).
101. See Estate of Thornton v. Caldor, Inc., 472 U.S. 703, 710 (1985); see also Texas Monthly, 489 U.S. at 18 n.8 (plurality opinion) (noting permissible accommodations did not impose substantial burdens on nonbeneficiaries); Amos, 483 U.S. at 337 n.15 (noting that the legislative accommodation impinged on third parties, but distinguishing this burden from the one held

1010
Constitutionality of RLUIPA

_Estate of Thornton v. Caldor, Inc._\(^{102}\) held that a statute violates the Establishment Clause when it imposes, without exception, an “absolute duty” on a non-beneficiary of the legislative accommodation.\(^{103}\) In _Caldor_, the Connecticut legislature revised its Sunday-closing law to permit certain classes of businesses to remain open on Sundays.\(^{104}\) The legislature also enacted a “Sabbath statute” prohibiting people from working more than six days in any calendar week and guaranteeing employees the right not to work on the Sabbath of their religious faith.\(^{105}\) The Sabbath statute forced employers to conform their business practices to the particular religious practices of their Sabbath-observing employees.\(^{106}\) The Court held that the Sabbath statute did not take any account of the significant burdens imposed on others, and this “unyielding weighting” of the Sabbath-observing employees’ interests over all other interests, including those of the employers and the non-Sabbath observing employees, advanced religion.\(^{107}\)

The Court in _Amos_ reaffirmed _Caldor_.\(^{108}\) However, in _Amos_, the Court held that § 702 of Title VII did not impermissibly burden third parties.\(^{109}\) While a religious organization might burden an employee by exercising its right under § 702, discriminating against the employee on the basis of religion would be permissible.\(^{110}\) The Court distinguished the burden under § 702 from the burden at issue in _Caldor_ on the grounds that the religious organization, not the state, imposed the § 702 burden.\(^{111}\)

C. Legislative Accommodations Must Cover an Appropriate Scope

Legislative accommodations benefiting a wide variety of groups, both religious and secular, are more likely to withstand an Establishment Clause attack than accommodations benefiting only religious entities or impermissible in _Caldor_).

---

103. Id. at 708–09.
104. Id. at 705 n.2.
105. Id.
106. Id. at 709.
107. Id. at 710.
109. Id.
110. Id.
111. Id.
individuals.\textsuperscript{112} In \textit{Walz v. Tax Commission of New York},\textsuperscript{113} the U.S. Supreme Court upheld a state property tax exemption that benefited religious organizations,\textsuperscript{114} as well as charitable and non-profit entities.\textsuperscript{115} The Court reasoned that the exemption was valid, even though it benefited religious organizations, because it did not single out a particular church, religious group, or even churches as a whole.\textsuperscript{116} Rather, the exemption recognized that entities fostering the community’s “moral or mental improvement”\textsuperscript{117} should not be inhibited in their activities by the chance that they would lose their properties for failure to pay property taxes.\textsuperscript{118}

In contrast to the broad tax exemption at issue in \textit{Walz}, the tax exemption in \textit{Texas Monthly} applied exclusively to religious periodicals and books.\textsuperscript{119} The \textit{Texas Monthly} Court held that the exemption violated the Establishment Clause.\textsuperscript{120} Although the majority determined that the tax exemption was an impermissible statutory preference for the dissemination of religious ideas,\textsuperscript{121} it did not agree on a rationale.\textsuperscript{122} A plurality of the Court took the most restrictive view of legislative

\textsuperscript{112} Compare \textit{Walz v. Tax Comm'n}, 397 U.S. 664, 672–73 (1970) (noting that the exemption was provided to “all houses of religious worship within a broad class of property owned by nonprofit, quasi-public corporations” that fostered the “moral or mental improvement” of the community), with \textit{Texas Monthly, Inc. v. Bullock}, 489 U.S. 1, 5 (1989) (plurality opinion) (holding that exemption “confined exclusively to publications advancing the tenets of a religious faith” ran afoul of the Establishment Clause).

\textsuperscript{113} 397 U.S. 664 (1970).

\textsuperscript{114} Id. at 680.

\textsuperscript{115} Id. at 673.

\textsuperscript{116} Id.

\textsuperscript{117} Id. at 672.

\textsuperscript{118} Id.


\textsuperscript{120} Id.

\textsuperscript{121} Id. at 28 (Blackmun, J., concurring). Under the “narrowest grounds” doctrine in \textit{Marks v. United States}, 430 U.S. 188, 193 (1977), the holding of the Court may be viewed as that position taken by the justices who concur in the judgments on the narrowest grounds. \textit{Id.} (citing \textit{Gregg v. Georgia}, 428 U.S. 153, 169 n.15 (1976)). The plurality in \textit{Texas Monthly}, consisting of Justices Brennan, Marshall, and Stevens, took the most restrictive view of permissible legislative accommodations. \textit{See Texas Monthly}, 489 U.S. at 10–12 (plurality opinion). The concurring opinion, written by Justice Blackmun and joined by Justice O’Connor, did not adopt the strict view of the plurality, but found that the statutory preference for religious ideas was impermissible under the Establishment Clause. \textit{See id.} at 26–28 (Blackmun, J., concurring). Because the plurality would concur with that view, \textit{id.} at 15, this would be the “narrowest grounds” and thus the holding in \textit{Texas Monthly}.

\textsuperscript{122} See \textit{Texas Monthly}, 489 U.S. at 5.
Constitutionality of RLUIPA

accommodations and found that an accommodation would be permissible only when its scope encompassed a broad class of both religious and non-religious entities. To withstand a challenge under the Establishment Clause, how broad the class required for inclusion in the exemption must be depends on the secular purpose of the exemption. An under-inclusion of groups benefiting from the accommodation would constitute a government endorsement offensive to the Establishment Clause. The plurality determined that every tax exemption constitutes a subsidy that affects nonqualifying taxpayers, forcing them to become indirect donors, and the Texas sales tax exemption lacked a secular objective that would justify this preferential benefit.

Nevertheless, the plurality acknowledged two exceptions where legislative accommodations for exclusively religious entities are permissible. First, the accommodation would be permissible when it would not impose substantial burdens on non-beneficiaries while allowing others to act according to their beliefs. Second, it would be permissible when the Free Exercise Clause required the accommodation.

Although the Texas Monthly tax exemption lacked sufficient breadth to pass scrutiny under the Establishment Clause, the Court has on other occasions upheld legislation benefiting religious entities exclusively on the grounds that the legislation has valid, secular purposes. For example, the Gillette Court upheld the conscientious

123. Id. at 11 (plurality opinion).
124. Id. at 15–16 (plurality opinion).
125. See id. at 16 (plurality opinion).
126. Id. at 14 (plurality opinion).
127. Id. at 17 (plurality opinion). However, even if the sales tax violated a religious group’s religious tenets, the Texas Monthly plurality was still hesitant to permit the accommodation when the Free Exercise Clause would not require it. Id. at 19 (plurality opinion). Regardless, the Smith Court later confirmed that there is an area of legislative discretion between what the Free Exercise Clause requires and what the Establishment Clause forbids. See Employment Div. v. Smith, 494 U.S. 872, 890 (1990). The Court also recently affirmed this in Locke v. Davey, 124 S. Ct. 1307, 1311 (2004).
128. Texas Monthly, 489 U.S at 18 n.8 (plurality opinion).
129. Id.
130. Id.
131. Id. at 14 (plurality opinion).
132. See, e.g., Corp. of the Presiding Bishop v. Amos, 483 U.S. 327, 335 (1987) (holding that alleviating governmental interference with the ability of religious organizations to define and carry out their mission is a valid secular purpose); Gillette v. United States, 401 U.S. 437, 453 (1971)
objector exemption, noting one of its valid secular purposes included an attempt to accommodate free exercise values and avoid conflicts with an individual’s religious conscience. Also, the Amos Court upheld § 702 of the CRA, even though it only benefits religious entities, because § 702 prevents a government regulation from burdening the exercise of religion. The Court noted that it had never indicated that statutes giving special consideration to religious groups were per se invalid, and when the government acted with the proper purpose of lifting a government regulation that burdened the exercise of religion, there was no need for a corresponding secular benefit.

In sum, the U.S. Supreme Court’s legislative accommodation precedent reveals three factors that determine whether an accommodation advances religion: the severity of the government-imposed burden on religion that the accommodation seeks to relieve, the burden that the accommodation places on third parties, and the scope of the accommodation. The severity of the government-imposed burden that the accommodation seeks to relieve must be significant, the accommodation must not substantially burden third parties, and the accommodation must not substantially burden third parties, and

133. Gillette, 401 U.S. at 453.
134. See Amos, 483 U.S. at 335–36.
135. Id.
136. Texas Monthly, 489 U.S. at 18 n.8 (plurality opinion); Amos, 483 U.S. at 338; Gillette, 401 U.S. at 445.
137. Texas Monthly, 489 U.S. at 18 n.8 (plurality opinion); Amos, 483 U.S. at 337 n.15; Estate of Thornton v. Caldor, 472 U.S. 703, 709 (1985).
138. See Texas Monthly, 489 U.S. at 11–17 (plurality opinion) (finding legislative accommodations permissible only, with a few exceptions, when they also benefit nonreligious groups, and requiring the accommodation’s purpose to determine the breadth of nonreligious groups that must be included); Amos, 483 U.S. at 338 (noting that when government acted with proper purpose of lifting burden on religious exercise, the action did not need to include secular entities); Gillette, 401 U.S. at 452–53 (noting the government’s valid purpose for limiting the scope of the accommodation based on pragmatic reasons as well as respect for the principle of supremacy of conscience); Walz v. Tax Comm’n, 397 U.S. 664, 672–73 (1970) (finding valid legislative purpose to limit accommodation to entities that foster “moral or mental improvement” of the community).
139. See Texas Monthly, 489 U.S. at 18 n.8 (plurality opinion) (finding “legislative exemptions . . . that were designed to relieve government intrusions that might significantly deter adherents of a particular faith from conduct protected by the Free Exercise Clause” permissible); Amos, 483 U.S. at 338 (finding it proper for a government regulation to lift burdens on religious exercise); Gillette, 401 U.S. at 445 (noting that the legislative accommodation was intended to prevent placing a burden on religious conscientious objectors to choose between religious conscience and suffering penalties).
140. See Texas Monthly, 489 U.S. at 18 n.8 (plurality opinion) (finding legislative accommodation that did not substantially burden third parties to be permissible); Amos, 483 U.S. at
Constitutionality of RLUIPA

the scope of the accommodation should be as broad as the secular purpose would require.  

III. CIRCUIT COURTS DISAGREE ABOUT WHETHER RLUIPA VIOLATES THE ESTABLISHMENT CLAUSE

The U.S. Courts of Appeals that have considered RLUIPA’s constitutionality disagree about whether the primary effect of the Act is to advance religion. The Fourth, Seventh, and Ninth Circuits have held that the primary effect of the Act is not to advance religion, while the Sixth Circuit has held that it is. In applying the Lemon analysis to determine RLUIPA’s primary effect, the Fourth, Seventh, and Ninth Circuits found that the Act lifts a burden on religious exercise. The Fourth Circuit concluded that the limitation of RLUIPA’s benefits to religious rights does not render the Act unconstitutional. Conversely, the Sixth Circuit held that the Act’s primary effect is to advance religion because it does not relieve a burden on religious exercise, it burdens third parties, and it is impermissibly limited in scope to religious entities only.

337 n.15 (noting another legislative accommodation that burdened third parties); Caldor, 472 U.S. at 709 (finding the legislative accommodation failed to consider the interests and convenience of third parties).

141. See Texas Monthly, 489 U.S. at 11–17 (plurality opinion) (finding legislative accommodations permissible only, with a few exceptions, when they also benefit nonreligious groups, and determining that the breadth of nonreligious groups that must be included in the accommodation depends on the accommodation’s purpose); Amos, 483 U.S. at 338 (finding legislative accommodations that exclusively benefit religious entities are not per se invalid and are permissible when lifting burdens on religion); Gillette, 401 U.S. at 452–53 (noting that the legislative accommodation’s purpose was not designed to favor any sect or religion); Walz, 397 U.S. at 672–73 (noting that the legislative accommodation did not single out particular religious entities or religion in general, where the purpose was to foster society’s mental or moral improvement).

142. Compare Madison v. Riter, 355 F.3d 310, 318 (4th Cir. 2003), and Charles v. Verhagen, 348 F.3d 601, 611 (7th Cir. 2003), and Mayweathers v. Newland, 314 F.3d 1062, 1068 (9th Cir. 2002), with Cutter v. Wilkinson, 349 F.3d 257, 267 (6th Cir. 2003).

143. Madison, 355 F.3d at 312; Charles, 348 F.3d at 611; Mayweathers, 314 F.3d at 1069.

144. Cutter, 349 F.3d at 267.

145. See Madison, 355 F.3d at 318; Charles, 348 F.3d at 611; Mayweathers, 314 F.3d at 1069.

146. See Madison, 355 F.3d at 318.

147. See Cutter, 349 F.3d at 267.

148. See id. at 266–67.

149. See id. at 264–66.
A. Several Circuit Courts Have Held that RLUIPA’s Primary Effect Is Not To Advance Religion

In *Madison v. Riter*, 150 *Charles v. Verhagen*, 151 and *Mayweathers v. Newland*, 152 the Fourth, Seventh, and Ninth Circuits, respectively, held that RLUIPA does not violate the Establishment Clause because its primary effect is not to advance religion. 153 In upholding RLUIPA, the courts considered, to varying degrees, the three factors the U.S. Supreme Court has weighed in its legislative accommodation precedent. 154 These courts concluded that RLUIPA lifts a burden on religious exercise. 155 The Seventh Circuit noted that RLUIPA sought only to remove the most substantial state-imposed burdens, while giving states’ penological interests due consideration. 156 Only the Fourth Circuit, in *Madison*, considered whether RLUIPA burdens third parties. 157 Among the third parties that RLUIPA possibly affects, the court considered only the states, and concluded that the Act would not constitute a burden. 158 It also found that the Act is not a burden because it applies only if the states choose to accept federal correctional funds. 159

Lastly, each court considered whether it was permissible for RLUIPA to relieve burdens on religious exercise without also relieving burdens on secular constitutional rights. 160 The courts agreed that RLUIPA could benefit religious rights only, 161 but only the Fourth Circuit, in *Madison*, provided a detailed analysis of this point. 162 In particular, the *Madison* court addressed the district court’s theory that Congress must be neutral toward religious entities and is not neutral when it lifts limitations on

150. 355 F.3d 310 (4th Cir. 2003).
151. 348 F.3d 601 (7th Cir. 2003).
152. 314 F.3d 1062 (9th Cir. 2002).
153. *Madison*, 355 F.3d at 313; *Charles*, 348 F.3d at 611; *Mayweathers*, 314 F.3d at 1069.
154. See *Madison*, 355 F.3d at 318, 320–21; *Charles*, 348 F.3d at 611; *Mayweathers*, 314 F.3d at 1069.
155. See *Madison*, 355 F.3d at 318; *Charles*, 348 F.3d at 611; *Mayweathers*, 314 F.3d at 1069.
156. See *Charles*, 348 F.3d at 611.
158. See id. at 321.
159. See id.
160. See id. at 318; *Charles*, 348 F.3d at 611; *Mayweathers*, 314 F.3d at 1068–69.
161. See *Madison*, 355 F.3d at 318; *Charles*, 348 F.3d at 611; *Mayweathers*, 314 F.3d at 1068–69.
Constitutionality of RLUIPA

religious rights while ignoring all others. The district court supported
this rationale by reading *Amos* to turn on a congressional finding that
Title VII threatened religious exercise more than other rights. The
Fourth Circuit rejected this interpretation of *Amos*. It concluded that
requiring Congress to examine if or how any other fundamental rights
were similarly burdened in order to protect burdened religious rights
would conflict with U.S. Supreme Court precedent. Instead, the
Fourth Circuit read *Amos* to hold that Congress may lift governmental
burdens on religious exercise without also relieving burdens on secular
constitutional rights.

B. The Sixth Circuit Has Held that RLUIPA’s Primary Effect Is To
Advance Religion

Breaking away from the other circuits, the Sixth Circuit, in *Cutter v.
Wilkinson*, held that RLUIPA violates the Establishment Clause. The
court considered three U.S. Supreme Court legislative
accommodation factors. In determining that RLUIPA’s primary effect
is to advance religion, the court first concluded that RLUIPA does not
lift a burden on religious exercise. It stated that the *Turner*
rational-relationship standard provided protection for religious exercise
at the time Congress enacted RLUIPA. The Act, therefore, did not rectify a

---

164. See *Madison*, 240 F. Supp. 2d at 577 n.9.
166. See *id.*
167. *Id.*
168. 349 F.3d 257 (6th Cir. 2003). While the Sixth Circuit applied the *Lemon* test as its
framework for analysis, *id.* at 262, it also considered the principle of endorsement from *Lynch v.
349 F.3d at 264. The Sixth Circuit focused on the primary effect prong and found that RLUIPA
conveyed a message of endorsement of religion. *Id.* It noted the U.S. Supreme Court had considered
several factors in making a determination of endorsement, *id.*, but the Sixth Circuit found only two
factors to be the most relevant in its inquiry—first, whether the particular government action
benefited both secular and religious entities, and second, whether the action would induce religious
exercise, rather than only protect it. *Id.*
170. *Id.* at 263–64, 266–67.
171. *Id.* at 267.
172. *Id.* at 263. See also supra Part I.C (describing the *Turner* rational-relationship test).
congressionally-created burden.\textsuperscript{173} Thus, an exclusively religious exemption, such as § 702 of the CRA, which Congress passed to relieve a burden Title VII placed on religious organizations, would not be warranted.\textsuperscript{174} The \textit{Cutter} court also addressed RLUIPA’s burden on third parties.\textsuperscript{175} The court concluded that the Act, which required enhanced protection for the right to exercise religion relative to other constitutional rights,\textsuperscript{176} favored religious rights over secular rights and would induce non-religious prisoners to adopt a religion in order to obtain a preferred status in the prison community.\textsuperscript{177}

Finally, addressing RLUIPA’s scope, the court concluded that the Act was not analogous to \textit{Amos} because it singled out religious rights for protection.\textsuperscript{178} Unlike the Fourth Circuit in \textit{Madison}, the Sixth Circuit in \textit{Cutter} interpreted \textit{Amos} to allow Congress to favor religious exercise rights over secular rights only when religious exercise rights are at a greater risk of deprivation.\textsuperscript{179} The \textit{Cutter} court noted that there was no evidence that prisoners’ religious exercise rights were threatened more than their secular rights.\textsuperscript{180} Furthermore, the court reasoned that any indication from \textit{Amos} that Congress could benefit only religious exercise rights could be explained.\textsuperscript{181} It interpreted the exemption in \textit{Amos} as necessary to avoid entangling the government in church operations to monitor their compliance with Title VII.\textsuperscript{182} It concluded that the \textit{Amos} accommodation was a narrowly tailored solution to the potential entanglement problem.\textsuperscript{183} Section 702 was necessary to avoid an Establishment Clause violation, whereas RLUIPA applied to many generally applicable prison regulations.\textsuperscript{184} The court reasoned that because the Free Exercise Clause did not require the enactment of

\begin{thebibliography}{184}
\bibitem{173} \textit{Cutter}, 349 F.3d at 267.
\bibitem{174} \textit{Id.}
\bibitem{175} See \textit{id.} at 266–67.
\bibitem{176} \textit{Id.} at 264–65.
\bibitem{177} \textit{Id.} at 266–67.
\bibitem{178} \textit{Id.} at 263–64. Although the court analyzed \textit{Amos} under the secular purpose prong of \textit{Lemon}, \textit{Amos} dealt with the scope factor in the primary effect prong. See Corp. of the Presiding Bishop v. Amos, 483 U.S. 327, 336–38 (1987).
\bibitem{179} See \textit{Cutter}, 349 F.3d at 265.
\bibitem{180} \textit{Id.} at 265–66.
\bibitem{181} \textit{Id.} at 268.
\bibitem{182} \textit{Id.} at 267.
\bibitem{183} \textit{Id.} at 268.
\bibitem{184} \textit{Id.}
\end{thebibliography}
Constitutionality of RLUIPA

RLUIPA, and there was no evidence that religious rights were at a greater risk of deprivation than other fundamental rights, RLUIPA impermissibly provided heightened protection only for religious prisoners.185

IV. AS A LEGISLATIVE ACCOMMODATION, RLUIPA DOES NOT VIOLATE THE ESTABLISHMENT CLAUSE

RLUIPA is a constitutional legislative accommodation under the Establishment Clause because it does not have the impermissible effect of advancing religion.186 U.S. Supreme Court precedent provides three relevant factors for courts to consider in making this determination.187 Applying these factors to RLUIPA demonstrates that the Act does not have the primary effect of advancing religion and is therefore constitutional under the Establishment Clause.188 First, RLUIPA lifts “substantial” burdens on a prisoner’s exercise of religion.189 Second, RLUIPA does not impermissibly burden third parties.190 Third, while RLUIPA protects only prisoners’ exercise of religion, the scope of the accommodation is appropriate given its secular purpose.191

A. RLUIPA Lifts Substantial Burdens on Prisoners’ Exercise of Religion

RLUIPA relieves “substantial” government-imposed burdens on prisoners’ religious exercise that do not further, by the least restrictive means, a compelling government interest.192 Thus, by its terms, RLUIPA brings itself within the rule established by the U.S. Supreme Court in Gillette and Amos, as well as the plurality in Texas Monthly: legislative accommodations may relieve only those “significant” burdens that

185. See id. at 265–66.
186. The secular purpose and entanglement prongs of the Lemon test as applied to RLUIPA are beyond the scope of this Comment.
188. See infra Part IV.A–C.
190. See infra Part IV.B.
191. See infra Part IV.C.
threaten “severe encroachment on protected religious freedoms.”

Prison regulations that prohibit or penalize inmates for following their religious beliefs constitute a significant burden. Prison regulations restricting religious exercise require prisoners to make the Hobson’s choice between acting in a manner offensive to their consciences and suffering the consequences of non-compliance. For example, Muslim and Jewish prisoners are sometimes served pork or forced to go hungry, though one Muslim minister declared that even “if our lives depend on it . . . we can’t eat pork.” Thus, there is a significant burden when prisoners must choose between obeying their religious consciences and complying with prison regulations.

All four circuit courts that have considered RLUIPA’s constitutionality incorrectly applied the severity of burden factor. Although the Fourth, Seventh, and Ninth Circuits correctly recognized that RLUIPA lifts a burden on prisoners’ religious exercise, the courts failed to analyze the severity of burden required for a permissible accommodation. The courts’ rulings would permit the accommodation of any burden on religion. However, if that were the case, the sales tax exemption in Texas Monthly would have been permissible.

The Sixth Circuit incorrectly concluded that prisoners’ religious exercise is not burdened because Congress did not impose an “affirmative” burden on religious exercise and enact RLUIPA as a

---

194. See Gillette, 401 U.S. at 445.
195. See id.
196. See, e.g., Hunafa v. Murphy, 907 F.2d 46, 48 (7th Cir. 1990) (finding genuine issue as to whether denial of a Muslim prisoner’s request for removal of pork from his meal tray would meet the rational-relationship standard); see also United States v. Huss, 394 F. Supp. 752, 753 (S.D.N.Y. 1975) (upholding the prison’s denial of a Jewish prisoner’s request for kosher food when applying a “reasonableness” standard), vacated on other grounds, 520 F.2d 598, 600 (2d Cir. 1975).
198. See Madison v. Riter, 355 F.3d 310, 318 (4th Cir. 2003); Charles v. Verhagen, 348 F.3d 601, 611 (7th Cir. 2003); Mayweathers v. Newland, 314 F.3d 1062, 1069 (9th Cir. 2002).
199. See Madison, 355 F.3d at 317–19; Charles, 348 F.3d at 609–11; Mayweathers, 314 F.3d at 1068–69.
200. See Madison, 355 F.3d at 317–19; Charles, 348 F.3d at 609–11; Mayweathers, 314 F.3d at 1068–69.
201. See Texas Monthly, Inc. v. Bullock, 489 U.S. 1, 18 n.8 (1989) (plurality opinion).
Constitutionality of RLUIPA

solution to the burden. In essence, the court distinguished a burden imposed by Congress itself, which may be lifted by a legislative accommodation enacted by Congress, from the burden imposed by rules or actions of any federal, state or local prison administration, which may not be lifted by a legislative accommodation enacted by Congress. Not only has the U.S. Supreme Court not made this distinction, but such a distinction does not make any substantive difference in the case of RLUIPA. In Amos, the Court stated that the government may act to lift a regulation that burdens religious exercise; however, the Court did not base its conclusion on the premise that Congress could only lift a burden it had directly imposed on religion. Furthermore, the Act applies only when the prison officials voluntarily accept federal funds or the burden on religious exercise affects interstate commerce. Therefore, by weighing the costs and benefits of adopting RLUIPA, prison officials may choose whether or not to adopt its requirements.

Second, the Sixth Circuit concluded that, in the absence of evidence that religious exercise had a “greater risk of deprivation,” RLUIPA had to protect “other fundamental rights” in order to be a permissible accommodation. The Sixth Circuit interpreted the severity of burden factor to require that the burden be more egregious or the right more likely to be deprived than other constitutional rights to warrant the accommodation. This stands contrary to Amos, which did not require the simultaneous protection of other fundamental rights for the accommodation to be permissible. The Sixth Circuit therefore was

203. See id.
205. Id.
206. See id.
209. While RLUIPA may apply to a prison that does not receive federal funding through the commerce provision in § 2000cc-1(b)(2), burdens on religious exercise or the removal of those burdens are unlikely to “affect commerce with foreign nations, among the several States, or with Indian tribes.” While the U.S. Supreme Court has interpreted the Commerce Clause to be quite expansive, see CHEMERINSKY, supra note 26, § 3.3.4, in recent years the Court has limited the scope of Congress’s commerce power. Id. § 3.3.5. However, the Commerce Clause is beyond the scope of this Comment.
211. Id.
incorrect because it found no burden on religious exercise when the burden is not directly imposed by Congress.\textsuperscript{213} Effectively, the Sixth Circuit required that there be greater deprivation of religious exercise relative to other fundamental rights before Congress could institute greater protection for religious exercise.\textsuperscript{214}

\section{RLUIPA Does Not Impermissibly Burden Third Parties}

While RLUIPA may inconvenience states by requiring them to comply with its conditions, this inconvenience does not qualify as a “burden” under U.S. Supreme Court precedent.\textsuperscript{215} As the \textit{Amos} Court indicated, to qualify as a burden for Establishment Clause purposes, the government itself must impose the limitation.\textsuperscript{216} In \textit{Amos}, the Court noted that Congress, through § 702 of the CRA, simply provided the opportunity for religious organizations to discriminate in choosing their employees; it was up to the organizations to burden their employees by exercising their rights under § 702.\textsuperscript{217} On the other hand, in \textit{Caldor}, the state directly burdened employers by passing a law requiring that they let their Sabbath-observing employees choose their day off.\textsuperscript{218} Under RLUIPA, the federal government does not directly burden states because RLUIPA only applies to the states if they accept federal funding for their prisons or the substantial burdens on religion affect interstate commerce.\textsuperscript{219} States may refuse to adopt RLUIPA’s conditions by forgoing federal funds.\textsuperscript{220}

RLUIPA also does not burden non-religious prisoners because RLUIPA does not lessen the protection afforded prisoners’ constitutional rights under the \textit{Turner} rational-relationship standard.\textsuperscript{221} If RLUIPA forced non-religious prisoners to give up their rights or to expend their own resources to help accommodate religious prisoners’ exercise of religion, its constitutionality could be questioned.\textsuperscript{222} However,
Constitutionality of RLUIPA

RLUIPA’s text does not explicitly require or even hint at such a possibility.\(^{223}\) Instead, the Act merely requires that regulations restricting religious exercise undergo greater scrutiny than regulations restricting other rights.\(^{224}\) Furthermore, the Act is not an absolute prohibition or demand as in *Caldor*.\(^{225}\) It requires only that prisons impose the least restrictive alternative,\(^ {226}\) which would permit prison officials to take into account the interests and rights of the non-religious prisoners.

The Seventh and Ninth Circuits did not consider RLUIPA’s burden on third parties,\(^ {227}\) and the Fourth Circuit did not consider RLUIPA’s burden on the prisoners who do not perform religious exercises.\(^ {228}\) The Sixth Circuit contemplated the possibility that RLUIPA would burden non-religious prisoners by causing them to resent their religious fellow inmates.\(^ {229}\) Although it is possible that non-religious prisoners may grow to resent their neighbors’ enhanced right to exercise their religion,\(^ {230}\) the U.S. Supreme Court has not addressed or considered resentment as a burden in accommodation cases. For instance, in *Gillette*, the Court did not raise the objection that people drafted to serve in the military might resent those left at home because their religion renounced all war.\(^ {231}\) Thus, while non-religious prisoners may resent the different treatment religious exercise receives, such resentment does not constitute the “unyielding weighting” of religious prisoners’ interests prohibited by *Caldor*.\(^ {232}\)

C. RLUIPA’s Scope Is Broad Enough, Given Its Secular Purpose

Although RLUIPA benefits only those prisoners who perform
religious exercises, the Act’s secular purpose of protecting religious liberty justifies its less-inclusive scope.\textsuperscript{233} Both the \textit{Gillette} and \textit{Amos} Courts upheld laws exclusively benefiting religious entities because the legislative intent to protect religious conscience and religious autonomy made it sensible and intuitive to protect only religious individuals or organizations.\textsuperscript{234} On the other hand, the \textit{Walz} Court recognized that when the government intended merely to foster society’s moral and mental improvement, the legislative accommodation appropriately covered religious, charitable, and quasi-public entities.\textsuperscript{235} While the \textit{Texas Monthly} Court did not reach a majority consensus on how inclusive or encompassing a tax exemption must be to be constitutional,\textsuperscript{236} the plurality described its view as requiring the accommodation to be as broad as its secular purpose would require,\textsuperscript{237} and also affirmed \textit{Amos}.\textsuperscript{238} While the plurality opinion is not binding precedent, its view is the strictest view of permissible legislative accommodations.\textsuperscript{239} Thus, if a legislative accommodation can meet the \textit{Texas Monthly} plurality’s strict standard, it would meet all other U.S. Supreme Court standards.

RLUIPA’s scope reaches all institutionalized persons exercising religion.\textsuperscript{240} Also, it does not, and need not, protect the exercise of “all other constitutional rights,”\textsuperscript{241} for the secular purpose of RLUIPA is “to protect religious liberty.”\textsuperscript{242} It is broad enough to cover all religious exercise, including all sects and minority religions.\textsuperscript{243} Non-religious individuals or entities thus would not need this protection, and including

\begin{footnotesize}
\begin{enumerate}
\item \textsuperscript{233} See, e.g., Corp. of the Presiding Bishop v. Amos, 483 U.S. 327, 338 (1987) (finding no reason to include secular entities in accommodation when purpose is to lift burden from religion).
\item \textsuperscript{234} See \textit{Amos}, 483 U.S. at 338; \textit{Gillette}, 401 U.S. at 452–53.
\item \textsuperscript{235} \textit{Walz} v. Tax Comm’n, 397 U.S. 664, 672–73 (1970).
\item \textsuperscript{236} See \textit{Texas Monthly, Inc. v. Bullock}, 489 U.S. 1, 16 (1989) (plurality opinion); \textit{id.} at 27–28 (Blackmun, J., concurring).
\item \textsuperscript{237} \textit{id.} at 15–16 (plurality opinion).
\item \textsuperscript{238} \textit{id.} at 18 n.8 (plurality opinion).
\item \textsuperscript{239} The Court has not articulated a more restrictive view of legislative accommodations provided exclusively for religious individuals or rights. See Shawn P. Bailey, \textit{The Establishment Clause and the Religious Land Use and Institutionalized Persons Act of 2000}, 16 \textit{Regent U. L. Rev.} 53, 64 n.82 (2004).
\item \textsuperscript{240} 42 U.S.C. § 2000cc-1 (2000). This assumes that the prison either receives federal funds or the burden affects interstate commerce. \textit{See supra} note 209.
\item \textsuperscript{241} 42 U.S.C. § 2000cc-1 (2000).
\item \textsuperscript{243} See 42 U.S.C. §§ 2000cc-1, cc-5(7)(A).
\end{enumerate}
\end{footnotesize}
Constitutionality of RLUIPA

them in the scope of this accommodation would not further this secular purpose. For example, it would be entirely illogical for RLUIPA, a federal statute that seeks to further the permissible secular purpose of protecting religious liberty, to include in its scope the right of prisoners to access pornography.\(^{244}\) Because the purpose of RLUIPA is secular and entirely permissible,\(^{245}\) this purpose controls the scope of the accommodation required to maintain constitutionality.\(^{246}\)

While the Fourth, Seventh, and Ninth Circuits correctly noted that the scope of a legislative accommodation may be limited exclusively to religious entities,\(^{247}\) these courts did not consider the breadth of scope required for the accommodation to be constitutional. The breadth required would depend on the scope of its secular purpose.\(^{248}\) In certain circumstances, accommodation of exclusively religious entities would be invalid if it was not broad enough to comport with its secular purpose.\(^{249}\)

However, the Sixth Circuit held that RLUIPA is unconstitutional.\(^{250}\) The Court concluded that, absent evidence that prisoners’ religious exercise rights are at “a greater risk of deprivation” than their other secular rights, RLUIPA cannot provide greater protection of prisoners’ religious exercise rights than their other secular rights.\(^{251}\) This standard is contrary to the holdings in Amos and Gillette, where the Court affirmed Congress’s ability to single out religious rights for enhanced protection when doing so is necessary to relieve those rights from a significant burden.\(^{252}\) Besides being contrary to U.S. Supreme Court precedent, the Sixth Circuit’s standard ignores that the appropriate

\(^{244}\) See Madison v. Riter, 355 F.3d 310, 319 (4th Cir. 2003).
\(^{245}\) This Comment assumes that RLUIPA’s purpose is not at issue. See also Corp. of the Presiding Bishop v. Amos, 483 U.S. 327, 335 (1987) (concluding that lifting government-imposed burdens on free exercise of religion is a proper government purpose).
\(^{247}\) Madison, 355 F.3d at 318; Charles v. Verhagen, 348 F.3d 601, 610 (7th Cir. 2003); Mayweathers v. Newland, 314 F.3d 1062, 1069 (9th Cir. 2002).
\(^{248}\) Texas Monthly, 489 U.S. at 15–16 (plurality opinion); Amos, 483 U.S. at 338; Gillette, 401 U.S. at 450–52; Walz, 397 U.S. at 672–73.
\(^{249}\) See, e.g., Texas Monthly, 489 U.S. at 15–16 (plurality opinion) (noting that an exclusively religious exemption would not be broad enough even when it had a valid purpose of promoting reflection and discussion about questions of ultimate value).
\(^{250}\) Cutter v. Wilkinson, 349 F.3d 257, 265 (6th Cir. 2003).
\(^{251}\) See id.
\(^{252}\) See Amos, 483 U.S. at 339; Gillette, 401 U.S. at 445.
secular purpose of RLUIPA is “to protect religious liberty.”

The Sixth Circuit ruling has wide-reaching implications. It calls into question many legislative accommodations. Some examples include the federal exemption from compulsory military service for ordained ministers and divinity students, and some states’ recognition of a “clergy-penitent” privilege. Furthermore, this standard creates an impossibly high standard that legislative accommodations must meet in order to withstand constitutional attack. In the Sixth Circuit, any time a legislature wishes to relieve an egregious burden on religious entities or individuals, it must determine what other fundamental rights are at risk and to what degree they are at risk. This requirement would lead to “byzantine complexities” that would prevent government from protecting fundamental rights above the minimum level set by the Constitution.

V. CONCLUSION

The U.S. Supreme Court should hold that RLUIPA does not violate the Establishment Clause because its primary effect is not to advance religion. Although legislatures may accommodate religion beyond what the Free Exercise Clause requires without violating the Establishment Clause, the Court should explicitly apply the three factors discussed above when analyzing whether the primary effect of a legislative accommodation advances religion, and find RLUIPA constitutional. This analysis will provide guidance to lower courts and aid legislatures in discerning the limits of their discretion when attempting to

---

254. See Madison v. Riter, 355 F.3d 310, 320 (4th Cir. 2003) (listing several examples of common legislative accommodations where constitutionality would be uncertain under the Sixth Circuit’s rationale).
255. Id.
256. Id.
257. See id.
258. Id.
259. Id. After Smith, this would preclude any protection above the minimum required by the U.S. Constitution. See Employment Div. v. Smith, 494 U.S. 872, 889 (1990). Furthermore, the Smith ruling suggests that, perhaps as compensation for diluting protection under the Free Exercise Clause, the U.S. Supreme Court intended to affirm the legislature’s latitude and discretion in accommodating burdens on religion and, in effect, invited greater legislative protection. See id. at 872, 890.
260. See Locke v. Davey, 124 S. Ct. 1307, 1311 (2004); Smith, 494 U.S. at 890.
Constitutionality of RLUIPA

accommodate religion. RLUIPA does not advance religion because it lifts substantial burdens on prisoners’ religious exercise, does not impermissibly burden third parties, and has an appropriate scope in light of its secular purpose to protect religious liberty.
**WHISTLING IN THE DARK? CORPORATE FRAUD, WHISTLEBLOWERS, AND THE IMPLICATIONS OF THE SARBANES–OXLEY ACT FOR EMPLOYMENT LAW**

Miriam A. Cherry

Abstract: Passed in 2002 in the wake of the accounting scandals that resulted in billions of dollars of lost value to shareholders, the Sarbanes–Oxley Act has as its major goal the prevention of corporate corruption. This Article analyzes the impact of section 806, the portion of the Sarbanes–Oxley Act that provides protections for employees who report securities fraud, and describes the effect that Sarbanes–Oxley has on existing employment law. In addition, this Article contributes to the debate over the general effectiveness of the Sarbanes–Oxley Act, a topic of contention among both academics and press commentators. This Article argues that the Act does not go far enough to protect whistleblowers because employers do not need to specify procedures for acting upon tips that allege financial fraud. Also, employers most likely can send whistleblowing claims to arbitration, a forum that weakens the remedies available to employees. Finally, this Article provides a comprehensive survey of state whistleblowing laws and suggests changes to federal and state law to fill the gaps that remain after Sarbanes–Oxley.

“Had it ever been a real company? Or had Enron been, from the very beginning, just a brilliant illusion?”

I. THE STATUS QUO: WHISTLEBLOWING LAW BEFORE SARBANES–OXLEY

A. Two Whistleblowers in Action: The Cases of WorldCom and Enron

1. Sherron Watkins Alerts Executives at Enron of Fraud

* Assistant Professor of Law, Cumberland School of Law, Samford University; B.A., 1996, Dartmouth College; J.D., 1999 Harvard Law School. This Article is possible because of the generous research support given by Samford University. I would like to acknowledge the law firm Berman, DeValerio, Pease, Tabacco, Burt & Pucillo, where I was encouraged to develop an expertise with securities fraud practice. In addition, I thank Sherron Watkins for sharing her experiences as a whistleblower, which were invaluable to this Article. Thanks to Edward Berger, III, John L. Carroll, Brannon P. Denning, Jill E. Evans, Leslie Griffin, Cameron S. Matheson, Michael T. Matraia, Angela Onwuechi-Willig, Jaimi A. Reisz, Robert L. Rogers, Kenneth M. Rosen, William G. Ross, Belle H. Stoddard, Stephen J. Ware, and the editors of the *Washington Law Review* for their helpful comments, suggestions, and encouragement.

2. Cynthia Cooper Uncovers WorldCom’s Massive Accounting Fraud .......................................................1039

B. Legal Status Quo: The Vagaries of Whistleblowing Law........................................................................1042
   1. State Public Policy Decisions and Whistleblower Statutes..........................................................1042
   2. Federal Whistleblower Statutes ..................................................................................................1049
   3. Sociological Approaches to Whistleblowing ...........................................................................1051

II. SARBANES–OXLEY ACT CHANGES TO SECURITIES LAW AND EMPLOYMENT LAW .....................1055
   A. General Provisions .................................................................1056
      1. Research Analysts and Conflicts of Interest ...............1057
      2. Increased Criminal Penalties for Securities Fraud .....1058
      3. Standardization of Accounting Practices ..................1059
      4. Corporate Governance .................................................1060
      5. Securities Fraud Litigation ..........................................1062
   B. Whistleblowing Provisions ....................................................1063
      1. Section 806: Federalizing the Law of Whistleblowing .........................................................1064
      2. Section 1107: Criminal Penalties for Shooting the Messenger ..............................................1067
      3. Section 301: Handling Whistleblower Complaints ....1069

III. THE EFFECTIVENESS OF SARBANES–OXLEY ..............1069
   A. Sarbanes–Oxley Fails to Specify Procedures that Employers Must Follow when Addressing Employee Complaints ........................................................................1070
   B. Mandatory Arbitration and Whistleblowing Claims ..............................1075
      1. Arbitration and the U.S. Supreme Court: From Skeptic to Supporter ....................................1076
      2. Will Sarbanes–Oxley Whistleblowing Claims be Sent to Arbitration? .................................1079
   C. Conclusions and Suggested Law Reform ....................1083

APPENDIX A ...................................................................................1087
APPENDIX B ....................................................................................1121
Employee Whistleblowing After Sarbanes–Oxley

As the accounting scandals surrounding Enron and WorldCom dominated the headlines and business ethics became increasingly suspect, two whistleblowers became symbols of integrity to the American public. Indeed, Sherron Watkins and Cynthia Cooper were among “The Whistleblowers” named as Time magazine’s “Persons of the Year” for 2002. At significant risk to their careers, financial well-being, and mental health, Cooper and Watkins alerted high-level executives at their respective companies to accounting fraud. Unfortunately, most whistleblowers take all these risks when they report illegal activities occurring within their organizations. The magnitude of these recent frauds is startling and, unfortunately, appears to be indicative of a widespread problem. One study by a prominent accounting firm estimated that companies lose twenty percent of every dollar earned due to some form of workplace fraud.

In response to the corporate scandals of 2002, Congress enacted the Sarbanes–Oxley Act (the Act) to prevent future corporate corruption and securities fraud. The Act contains a provision, § 806, that aims to

2. Richard Lacayo & Amanda Ripley, Persons of the Year, Time, Dec. 30, 2002, at 30. Sherron Watkins remarked that she had several similarities with the two other whistleblowers: all three are women, all three were working as primary wage-earners, and all three grew up in small towns. Interview with Sherron Watkins in Birmingham, Ala. (Feb. 18, 2004).
4. According to a study of eighty-four whistleblowers conducted in the early 1990s, “82% experienced harassment after blowing the whistle, 60% were fired, 17% lost their homes, and 10% admitted to attempted suicide.” David Culp, Whistleblowers: Corporate Anarchists or Heroes? Towards a Judicial Perspective, 13 Hofstra Lab. & Emp. L.J. 109, 113 (1995). For a fuller description of the risks accompanying whistleblowing, see infra Part I.B.3.
6. See id. The study was based on a telephone survey of 617 workers. Id.; see also Stephanie Armour, More Companies Urge Workers to Blow the Whistle, USA Today, Dec. 16, 2002, at B1 (reporting a trend toward whistleblowing as a method of curtailing workplace fraud). Of course, “fraud in the workplace” can mean a type of theft that is as trivial as making a personal copy at work or taking five extra minutes on a break. The type of fraud that appears to be most damaging and devastating is that perpetrated by individuals at the top of the organizational hierarchy—the magnitude of these frauds is greater because of the institutional trust given to leaders of an organization. See generally Gerard M. Zack, Fraud and Abuse in Nonprofit Organizations: A Guide to Prevention and Detection 20 (2003) (describing frauds and fraud-prevention techniques in the context of non-profit organizations).
protect whistleblowers such as Cooper and Watkins who report accounting fraud. Protect whistleblowers such as Cooper and Watkins who report accounting fraud. Most of the current scholarship on Sarbanes–Oxley focuses on the § 307 ethics rules that have been proposed to address a lawyer’s behavior when he or she is presented with knowledge of a client’s falsification of financial results (the so-called “noisy withdrawal” or “lawyer-as-whistleblower” provisions). This Article instead examines the provisions of Sarbanes–Oxley that cover all workers at publicly traded companies who “blow the whistle” on suspect accounting practices, whether that whistleblowing is done within the organization, to government agencies, or as part of a shareholder investigation.


10. Id.


Employee Whistleblowing After Sarbanes–Oxley

lawsuit. This Article discusses the few publicized cases that have been brought under § 806 to date. In discussing the whistleblower provisions, this Article uses an employment law perspective to examine the changes that Sarbanes–Oxley makes to existing state retaliatory discharge laws.

“Traditional” whistleblowing doctrine has long been the province of individual states, and as such, legal protection for whistleblowers has been largely decentralized and uneven. An initial examination of the whistleblowing provisions of the Act suggests that the new law ameliorates this traditional unevenness. However, a further and fuller analysis reveals that the lack of remedies provided in the Act results in a less effective scheme for encouraging whistleblowing than it would initially seem. Further, even after the enactment of Sarbanes–Oxley there are significant gaps in retaliatory discharge law. For example, while Sarbanes–Oxley does contain criminal provisions that cover retaliation against whistleblowers at all workplaces, plaintiffs have a private right of action under § 806 only if their employer is a publicly traded company. Therefore, it would appear that Sarbanes–Oxley incrementally changes both securities law and employment law, rather than radically overhauling either field. Accordingly, one might ask whether § 806 is substantive or merely “whistling in the dark.”

13. Recently, an administrative law judge ruled in favor of a chief financial officer who claims that he had been fired for raising various concerns regarding the issuer’s financial statements, including the chief executive officer’s insider trading. Welch v. Cardinal Bankshares Corp., 2003-SOX-15 (Dep’t of Labor Jan. 28, 2004), available at http://www.oalj.dol.gov/public/wblower/descn/03sox15c.htm (last visited Oct. 7, 2004). The interpretations that this particular case gives to provisions of the Act will be discussed throughout the Article. Another such whistleblowing case is being litigated in federal court in Texas. Murray v. TXU Corp., 279 F. Supp. 2d 799, 799 (N.D. Tex. 2003) (allowing suit to proceed despite defendant’s challenge to procedure and timing). The Murray case involves a high-level executive at an energy trading company who alleges that he was fired for raising internal concerns about earnings manipulation, lack of meaningful financial disclosures, and false and misleading statements by the company’s management. Id.; see Sudeep Reddy, TXU Denies Ex-VP’s Allegations: Utility Submits First Detailed Response to Whistleblower’s Suit, DALLAS MORNING NEWS, Sept. 17, 2003, at 2D. For discussion of a contrasting Sarbanes–Oxley whistleblower case that was sent to mandatory arbitration rather than to court, see this Article’s discussion of Boss v. Salomon Smith Barney, Inc., 263 F. Supp. 2d 684 (S.D.N.Y. 2003). See infra notes 362–376 and accompanying text.

17. The tension within the doctrine surrounding whistleblowing arises from the economic analysis that underlies so many areas of employment law—legal rules should encourage the socially optimal amount of whistleblowing activity, no more and no less. With too little whistleblowing activity, we
This Article examines whether, overall, Sarbanes–Oxley is a radical departure that will drastically change securities law and corporate governance, or whether it merely gives the appearance of change while leaving the status quo in place. As one commentator put it, Sarbanes–Oxley “is not major reform, but patches and codifications and further study. It is a restatement with the force of federal law.” Based on analysis of the whistleblowing provisions, this Article largely concurs and concludes that Sarbanes–Oxley is a compromise half-measure and not the true reform our securities laws need to respond to corporate fraud.

Part I of this Article examines whistleblower protections before the passage of Sarbanes–Oxley. It includes the stories of two recent whistleblowers on corporate fraud, Watkins and Cooper, and analyzes their legal situations prior to the passage of the Sarbanes–Oxley Act. This section also describes the status of state and federal protections for whistleblowers before the passage of the Act and the sociological implications of whistleblowing. The next section, Part II, describes and analyzes the various provisions of the Sarbanes–Oxley Act, including an in-depth analysis of the Act and regulations pertaining to whistleblowing. Part III considers the overall effectiveness of the Act. Ultimately, this Article concludes that Congress should have gone further to protect whistleblowers in the Sarbanes–Oxley Act and are left with sycophants who will agree to go on midnight toxic-dumping sprees, tell the Jeff Skillings of the world what they want to hear about ever-more-fraudulent related-party transactions, and allow polluted water into healing baths. On the other hand, with too much whistleblowing, we are left with organizations hobbled by internal strife. The same sort of tension between over- and under-reporting exists in other areas as well, and these same debates have gone on for years in terms of the reporting of racial and sexual harassment. See Richard A. Epstein, Forbidden Grounds: The Case Against Employment Discrimination Laws (1992).

18. Lawrence A. Cunningham, The Sarbanes–Oxley Yawn: Heavy Rhetoric, Light Reform (and it Just Might Work), 35 Conn. L. Rev. 915, 987 (2003) (emphasis omitted). In this analysis, I largely concur with Cunningham’s article, which describes the Sarbanes–Oxley Act as containing “sweeping punts and stunts,” but also as an Act that contains “potentially profound” provisions. Id. at 918–19. Cunningham states that “[i]ncremental provisions of the Act are best seen as patchwork responses to precise transgressions present in the popularized scandals—legislative action akin to the frequently maligned military strategist fighting the last war rather than planning for the next.” Id. At the same time, however, I disagree with Cunningham’s brief assessment of whistleblowing under Sarbanes–Oxley, as he seems to provide an overly optimistic view of the strength of state whistleblowing laws. See id. at 965–66. Another commentator remarked that, while “no panacea,” the Sarbanes–Oxley Act should be applauded for its steps to minimize conflicts of interest and for the publicity it is giving to business ethics. Robert Prentice, Enron: A Brief Behavioral Autopsy, 40 Am. Bus. L.J. 417, 442–43 (2003). Both of these points are important, Prentice argues, as they may affect behavior—not just rational behavior under the law and economics school of thought, but also behavior limited by bounded rationality and other such constraints. Id. at 443–44.
Employee Whistleblowing After Sarbanes–Oxley

legislators should amend various federal and state laws to fill in the gaps that make whistleblowers vulnerable to retaliation.

I. THE STATUS QUO: WHISTLEBLOWING LAW BEFORE SARBANES–OXLEY

A. Two Whistleblowers in Action: The Cases of WorldCom and Enron

Sherron Watkins and Cynthia Cooper played instrumental roles in exposing corporate fraud at their respective companies. While Watkins’s whistleblowing activity would not have been protected under Texas law prior to the passage of the Sarbanes–Oxley Act, Cooper’s whistleblowing would have been protected under Mississippi law. There is, however, no principled distinction that can be drawn between their two situations. Given the extent of the frauds and the problems that such accounting scandals created, these two situations justify a uniform federal system, such as that embodied by Sarbanes–Oxley.

1. Sherron Watkins Alerts Executives at Enron of Fraud

Sherron Watkins, trained as an accountant, had worked for Andrew Fastow, Enron’s chief financial officer, for nearly a decade in various capacities. During 2001 and 2002, Watkins, then a vice president at Enron, was charged with examining Enron’s books to find assets that the company could sell in response to the stock market decline. As she was assessing sale values, Watkins began finding “mystery assets” and “fuzzy off-the-books arrangements that seemed to be backed by nothing more than . . . deflated Enron stock.” When she asked questions, no one cared or seemed to be able to explain the arrangements.

19. See Austin v. HealthTrust, Inc., 967 S.W.2d 400, 403 (Tex. 1998); Winters v. Houston Chronicle Publ’g Co., 795 S.W.2d 723, 724 (Tex. 1990) (refusing to recognize “cause of action for private employees who are discharged for reporting illegal activities”); cf. Sabine Pilot Serv., Inc. v. Hauck, 687 S.W.2d 733, 734 (Tex. 1985) (creating “very narrow exception to the employment-at-will doctrine,” but requiring plaintiff “to prove by a preponderance of the evidence that his discharge was for no reason other than his refusal to perform an illegal act”).


21. See SWARTZ & WATKINS, supra note 1, at 269.

22. See id.


25. Id.
she learned, the more Watkins became “highly alarmed” because her “understanding as an accountant [was] that a company could never use its own stock to generate a gain or avoid a loss on its income statement.”26 However, Watkins felt that bringing her concerns either to Fastow or to Jeffrey Skilling, Enron’s then-chief executive officer, “would have been a job terminating move.”27

Concluding that “accounting should just not be that creative,” Watkins began looking for a job outside Enron.28 As she was in the process of looking for a position with another company, Skilling abruptly resigned, and Kenneth Lay, the chairman of the board, asked employees to drop suggestions in a comment box.29 Watkins composed a now-famous anonymous letter warning that the accounting for certain transactions had been too aggressive.30 The letter voiced her concern about improprieties and violations of accounting standards.31 Soon after

27. Id.
29. Morse & Bower, supra note 24, at 55; Interview with Sherron Watkins in Birmingham, Ala. (Feb. 18, 2004).
30. SWARTZ & WATKINS, supra note 1, at 275.
31. The letter, which has been widely reprinted, reads as follows:

Dear Mr. Lay,

Has Enron become a risky place to work? For those of us who didn’t get rich over the last few years, can we afford to stay?

Skilling’s abrupt departure will raise suspicions of accounting improprieties and valuation issues. Enron has been very aggressive in its accounting—most notably the Raptor transactions and the Condor vehicle. We do have valuation issues with our international assets and possibly some of our EES MTM positions.

The spotlight will be on us, the market just can’t accept that Skilling is leaving his dream job. I think that the valuation issues can be fixed and reported with other goodwill write-downs to occur in 2002. How do we fix the Raptor and Condor deals? They unwind in 2002 and 2003, we will have to pony up Enron stock and that won’t go unnoticed.

To the layman on the street, it will look like we recognized fund flows of $800 mm from merchant asset sales in 1999 by selling to a vehicle (Condor) that we capitalized with a promise of Enron stock in later years. Is that really funds flow or is it cash from equity issuance?

We have recognized over $550 million of fair value gains on stocks via our swaps with Raptor, much of that stock has declined significantly—Avici by 98%, from $178 mm to $5 mm, The New Power Co by 70%, from $20/share to $6/share. The value in the swaps won’t be there for Raptor, so once again Enron will issue stock to offset these losses. Raptor is an LJM entity. It sure looks to the layman on the street that we are hiding losses in a related company and will compensate that company with Enron stock in the future.

I am incredibly nervous that we will implode in a wave of accounting scandals. My 8 years of Enron work history will be worth nothing on my resume, the business world will consider the past successes as nothing but an elaborate accounting hoax. Skilling is resigning now for “personal reasons” but I think he wasn’t having fun, looked down the road and knew this stuff was unfixable and would rather abandon ship now than resign in shame in 2 years.
Employee Whistleblowing After Sarbanes–Oxley

writing the letter, Watkins arranged a meeting with Lay in which she expressed her concerns and was told that there would be an investigation into the points she raised.32

Months later, the Enron bankruptcy and fraud became well publicized. Indeed, Fastow and other executives had arranged for Enron to “hedge” with private partnerships, with many of the transactions backed only by Enron stock.33 Watkins was called to testify about this scheme before Congress.34 In the process of preparing for her testimony, Watkins reviewed an e-mail message from Enron’s outside counsel, Vinson & Elkins, which discussed the state of Texas whistleblowing law.35 This message read, in part, “[p]er your request . . . the following are some bullet thoughts on how to manage the case with the employee who made the sensitive report . . . . Texas law does not currently protect corporate whistle-blowers.”36 Watkins realized that the e-mail was in reference to her, as it was dated only two days after she met with Lay.37

The legal implication of this e-mail message is clear. Evidently, someone at a high level of the company wanted to fire Watkins because of her willingness to come forward and discuss the accounting improprieties. As Watkins put it, “I found out . . . that Ken Lay’s first action was not to look at my concerns [about fraudulent accounting] but

Is there a way our accounting guru’s [sic] can unwind these deals now? I have thought and thought about how to do this, but I keep bumping into one big problem—we booked the Condor and Raptor deals in 1999 and 2000, we enjoyed a wonderfully high stock price, many executives sold stock, we then try and reverse or fix the deals in 2001 and it’s a bit like robbing the bank in one year and trying to pay it back 2 years later. Nice try, but investors were hurt, they bought at $70 and $80/share looking for $120/share and now they’re at $38 or worse. We are under too much scrutiny and there are probably one or two disgruntled “redeployed” employees who know enough about the “funny” accounting to get us in trouble. What do we do? I know this question cannot be addressed in the all employee meeting, but can you give some assurances that you and Causey will sit down and take a good hard objective look at what is going to happen to Condor and Raptor in 2002 and 2003?

Id. at 361–62.
32. Id. at 287.
33. Michael L. Fox, To Tell or Not to Tell: Legal Ethics and Disclosure After Enron, 2002 COLUM. BUS. L. REV. 867, 870–80. In this instance, it was not hedging alone that was the problem. A “hedge” is often used to control the amount of risk to which an investor or an entity is subjected based on a downturn in the market. This explains the nature of hedge funds, which are countercyclical to the business cycle. The problem with Enron’s hedging was that it used its own assets. Not only does this violate various accounting rules, but instead of minimizing risk, it actually increases the amount of risk. See SWARTZ & WATKINS, supra note 1, at 228–34 (describing creation of Raptor entities for hedging purposes).
34. SWARTZ & WATKINS, supra note 1, at 352–53.
35. Morse & Bower, supra note 24, at 53.
36. Id.; see SWARTZ & WATKINS, supra note 1, at 291.
37. SWARTZ & WATKINS, supra note 1, at 291.
to see if they could dump me on the street.”38 In her view, Enron executives asked outside counsel to research the status of Texas whistleblower law so that Enron could fire her without employment law liability.39

Indeed, the e-mail embodied an accurate legal opinion because Texas law at that time would have allowed Lay to fire Watkins with no liability or legal consequence for Enron. In *Austin v. HealthTrust, Inc.*,40 the Supreme Court of Texas refused to recognize a retaliatory discharge cause of action based on an employee’s report of illegal activity at a private company (as opposed to a government employer).41 Before Enron was able to fire Watkins, however, the scandals about the fraudulent transactions came to light and Enron’s financial practices were subjected to scrutiny by its shareholders, Congress, and the American public.

The unraveling of the Enron saga continues.42 Currently, one Enron ex-treasurer, Ben Glisan, is in federal prison serving a five-year sentence that he received after pleading guilty to one count of conspiring to commit fraud.43 Fastow, who originally maintained his innocence,44 entered a guilty plea in exchange for a ten-year prison sentence and has now agreed to cooperate with law enforcement.45 Prosecutors recently filed a forty-two-count indictment against Skilling, who pleaded not guilty and is currently free on a $5 million bond.46 Former Chief

38. Sherron Watkins, Address at Samford University (Feb. 19, 2004).
39. Id.
40. 967 S.W.2d 400 (Tex. 1998).
41. Id. at 403 (refusing to recognize action for whistleblowing where nurse was allegedly retaliated against for reporting a co-worker who was endangering patients through substance abuse and illegal distribution of prescription drugs); see also Winters v. Houston Chronicle Pub’g Co., 795 S.W.2d 723, 724 (Tex. 1985) (refusing to recognize “cause of action for private employees who are discharged for reporting illegal activities”); Sabine Pilot Serv., Inc. v. Hauck, 687 S.W.2d 733, 734 (Tex. 1985) (creating “very narrow exception to the employment-at-will doctrine,” but requiring plaintiff “to prove by a preponderance of the evidence that his discharge was for no reason other than his refusal to perform an illegal act”).
44. Id.
Employee Whistleblowing After Sarbanes–Oxley

Accounting Officer Richard Causey is also awaiting trial.47 Meanwhile, a federal grand jury indicted three Merrill Lynch executives on fraud charges arising from a sham transaction that allowed Enron to record $12 million in earnings based on the ostensible purchase of Nigerian barges.48 Last, but certainly not least, Lay himself has been indicted by a grand jury and faces civil charges from the Securities and Exchange Commission (SEC) that could force him to disgorge his earnings from the sale of Enron stock and return those gains to investors.49

Enron, at one time the seventh largest publicly traded company on the New York Stock Exchange and a pioneer in new economy energy trading, now trades only as a penny stock.50 Enron’s stock certificates are currently sold for their historical and artistic value rather than for the worth of the company that the certificates represent.51 The multicolored sign “E” that stood outside Enron’s office, symbol of a once-proud corporation, is now better known as the “crooked ‘E,’” a symbol of corporate corruption.52

2. Cynthia Cooper Uncovers WorldCom’s Massive Accounting Fraud

In March of 2002, Cynthia Cooper, the head of internal auditing at WorldCom, heard a disturbing report about WorldCom’s accounting from an executive in the wireless division, who maintained that WorldCom used certain reserves to boost its revenue.53 When Cooper brought the matter to the attention of outside auditor Arthur Andersen, she was told that the accounting treatment was not a problem and could

47. Id.
51. A visit to the website eBay, http://www.ebay.com, on Oct. 7, 2004, confirmed that Enron stock certificates are worth more as art or as a symbol of a once high-flying company than they are as securities qua securities.
be ignored.\textsuperscript{54} When she raised the issue with WorldCom Chief Financial Officer Scott Sullivan, he became angry.\textsuperscript{55} Still concerned about the incident as well as other potential accounting irregularities, Cooper and her team of auditors secretly logged into the company’s computer system at night to check the accuracy of the accounts.\textsuperscript{56}

What Cooper and her employees found was shocking. In May of 2002, they discovered that WorldCom had treated operating costs as capital expenditures.\textsuperscript{57} While operating costs (such as salaries) are to be subtracted from income during the quarter, capital costs are major purchases (such as equipment or property) that can be spread out over time.\textsuperscript{58} By falsely reclassifying operating costs as capital expenditures in violation of accounting rules, WorldCom looked far more profitable in its financial reporting than it actually was.\textsuperscript{59}

That June, after her team turned up over $2 billion of suspect accounting entries, Cooper informed the board’s audit committee of her findings.\textsuperscript{60} After giving Sullivan a weekend to explain the accounting treatment, which he could not do, the audit committee fired him.\textsuperscript{61} Soon thereafter, WorldCom was forced to admit that it had inflated its profits by $3.9 billion.\textsuperscript{62} The truth was revealed in newspaper accounts, SEC filings, and congressional hearings: top management had directed several accountants to compensate for earnings short-falls at the end of each quarter when results were due by reclassifying operating costs as capital expenditures.\textsuperscript{63}

Two external investigations released in 2003 document the extent of

\begin{itemize}
\item \textsuperscript{54} Id.
\item \textsuperscript{55} Id.
\item \textsuperscript{56} Id.
\item \textsuperscript{57} Id.; see Lynne W. Jeter, \textit{Disconnected: Deceit and Betrayal at WorldCom} 168–73 (2003).
\item \textsuperscript{59} Id.
\item \textsuperscript{60} Pulliam & Solomon, supra note 53, at A6.
\item \textsuperscript{61} Id.
\item \textsuperscript{62} Andrew Hill, \textit{Whistleblowers’ Early Warning on WorldCom}, \textit{FIN. TIMES}, July 15, 2002, at 16.
\end{itemize}
Employee Whistleblowing After Sarbanes–Oxley

the fraud.64 These investigations reveal that WorldCom kept two sets of books and that there were numerous failures in the company’s corporate governance structure.65 In addition to the switches surrounding capital expenses and operating costs, the company also manipulated revenues, depreciation reserves, and even taxes.66 Sullivan and four other WorldCom executives and employees face criminal charges for orchestrating the $3.8 billion fraud.67 The Oklahoma Attorney General has indicted former Chief Executive Officer Bernard Ebbers, who was in charge of the company when the massive fraud took place and knew about the exaggerated numbers at the end of each quarter.68 Ebbers also may face federal charges.69

According to The Wall Street Journal, “[o]ne of the reports said while dozens of people knew about the fraud, it remained hidden from public view because employees were afraid to speak out.”70 Although Cooper has kept her position at WorldCom, reporting the fraud took a heavy personal toll.71 Even though Cooper did not perpetuate the fraud, many at the company blamed her for her discovery.72 The fraud led to massive layoffs and WorldCom’s subsequent bankruptcy.73

In contrast to Sherron Watkins, whose whistleblowing would have been unprotected in Texas,74 Mississippi state law likely would have protected Cooper.75 In a 1993 decision, McArn v. Allied Bruce-Terminix Co.,76 the Mississippi State Supreme Court created a public policy

64. See Rebecca Blumenstein & Susan Pulliam, WorldCom Fraud Was Widespread: Ebbers, Many Executives Conspired to Falsify Results in Late 1990s, Probes Find, WALL ST. J., June 10, 2003, at A3.
65. Id.
66. Id.
69. Brickey, supra note 42, at 358 n.4.
70. Blumenstein & Pulliam, supra note 64, at A3.
72. Id.
73. Id.
74. See Austin v. HealthTrust, Inc., 967 S.W.2d 400, 403 (Tex. 1998).
75. I have assumed for purposes of this Article that if, hypothetically, adverse employment action had been taken against Cooper, she would have brought an action under Mississippi whistleblower laws. WorldCom’s corporate headquarters were in Mississippi, and that is where Cooper worked during her entire tenure with WorldCom. See JETER, supra note 57, at 167–68.
76. 626 So. 2d 603 (Miss. 1993).
exception to the at-will employment rule for whistleblowers. Cooper and Watkins each had to make a choice about whether to confront wrongdoing and report fraud. Yet Watkins, facing this dilemma in Texas, could have been fired without legal recourse, whereas Mississippi law protected Cooper from that fate. These two instances highlight the inconsistencies in state law that existed prior to the passage of Sarbanes–Oxley.

B. Legal Status Quo: The Vagaries of Whistleblowing Law

This section summarizes state and federal whistleblowing laws prior to the Sarbanes–Oxley Act. Under traditional common law, most jurisdictions provided no legal redress for employees whose employers terminated them in retaliation for their whistleblowing. Increasingly, jurisdictions are moving from this legal regime to one that protects individuals who report wrongdoing at work. However, the law has developed somewhat haphazardly. As illustrated by the examples of Sherron Watkins and Cynthia Cooper, whether an employee can bring a successful case often depends on the state in which the employee blows the whistle.

1. State Public Policy Decisions and Whistleblower Statutes

Under the common law of various states, whistleblowing cases—sometimes classified as retaliatory discharge, discharge in contravention of public policy, or wrongful discharge—have received, and continue to receive, inconsistent treatment. An action for wrongful discharge is an exception to the traditional employment at-will doctrine. Under the at-will doctrine, an employer can hire and fire employees for a good

77. Id. at 607.
78. Compare Austin, 967 S.W.2d at 403, with McArn, 626 So. 2d at 607.
82. Id.
83. See, e.g., Sterling Drug, Inc. v. Oxford, 743 S.W.2d 380, 385 (Ark. 1988) ("[W]e hold that an at-will employee has a cause of action for wrongful discharge if he or she is fired in violation of a well-established public policy of the state. This is a limited exception to the employment-at-will doctrine.").
reason, any reason, or no reason at all; in parallel, employees can quit for a good reason, any reason, or no reason at all. As the common law developed, however, courts deemed certain reasons to be against public policy, such as firing an employee for refusing to give perjured testimony to a state investigatory committee or for serving as a juror.

From these initial limited public policy exceptions, courts in some jurisdictions began fashioning rudimentary protection for employees who reported illegal actions that occurred at work. Courts in early decisions reasoned that an employee should not be fired for refusing to violate a law or government regulation or for revealing such a violation. One detects in the cases a certain hesitancy by the courts to intrude into the “private” decision to discharge an employee. At the

84. Development of the at-will doctrine is usually attributed to a 125-year-old treatise by Horace G. Wood. Wood, supra note 79, at 272. For a useful discussion of the rule, as well as its benefits and drawbacks, see Paul C. Weiler, Governing the Workplace: The Future of Labor and Employment Law 56–63 (1990).

85. Indeed, the genesis of whistleblowing doctrine is often traced to a case about perjured testimony. In Petermann v. Int’l Bhd. of Teamsters, the court, after establishing that plaintiff had a cause of action for wrongful discharge when he was fired for refusing to commit perjury, stated: To hold that one’s continued employment could be made contingent upon his commission of a felonious act at the instance of his employer would be to encourage criminal conduct upon the part of both the employee and employer and would serve to contaminate the honest administration of public affairs. This is patently contrary to the public welfare.


86. See Nees v. Hocks, 536 P.2d 512, 516 (Or. 1975) (creating exception to the employment-at-will rule and allowing plaintiff to recover when she was dismissed for serving on a jury). The court also stated, “[i]f an employer were permitted with impunity to discharge an employee for fulfilling her obligation of jury duty, the jury system would be adversely affected. The will of the community would be thwarted.” Id.


88. See, e.g., Tameny v. Atl. Richfield Co., 610 P.2d 1330, 1335–37 (Cal. 1980) (holding that employee fired after refusing to participate in employer’s illegal scheme stated cause of action for wrongful discharge); Johnson, 433 N.W.2d at 228 (holding that employee engaged in protected activity in reporting corporate president’s conversion of company property to own use).


We are mindful that courts should not lightly intervene to impair the exercise of managerial discretion or to foment unwarranted litigation. We are, however, equally mindful that the myriad of employees without the bargaining power to command employment contracts for a definite term are entitled to a modicum of judicial protection when their conduct as good citizens is punished by their employers.


[T]he rule giving the employer the absolute right to discharge an at-will employee must be tempered by the further principle that where the employer’s motivation for the discharge contravenes some substantial public policy principle, then the employer may be liable to the employee for damages occasioned by the discharge.
same time, however, many courts overcame this hesitancy by focusing on the legal violations that took place. To allow employers to direct their employees to break the law, to fire employees if they refused to comply, and then to deny the discharged employee any kind of legal redress would be tantamount to encouraging employers to break the law. Instead of relying on public policy, other courts fashioned remedies for aggrieved employees by implying a contractual duty of good faith and fair dealing into the at-will employment relationship. Not all state courts, however, have recognized a cause of action for such a retaliatory discharge. In many jurisdictions, such as Texas, where Sherron Watkins would likely have had her case heard if Enron had fired her, courts have refused to fashion such a public policy exception to the at-will rule.

Although academics have soundly criticized the employment at-will rule, and debate over the rule’s wisdom continues, it is the default rule in all but one American jurisdiction. Currently, only Montana has abrogated the rule and adopted a for-cause termination statute. However, the trend among the other forty-nine states has been to temper the at-will rule somewhat. State legislatures, state courts, and Congress have increasingly encroached upon the at-will rule by providing some employees a limited level of job security based on public policy

---

90. See, e.g., Beasley v. Affiliated Hosp. Prods., 713 S.W.2d 557, 559 (Mo. Ct. App. 1986) (involving employee who refused to determine outcome of raffle, which would have violated state law).

91. See, e.g., Monge v. Beebe Rubber Co., 316 A.2d 549, 551 (N.H. 1974) (implying duty of good faith into contract of employment that was terminable at will).


Employee Whistleblowing After Sarbanes–Oxley

reasons. In response to the perceived harshness of the at-will rule and, in some instances, in response to judicial unwillingness to provide employees with a remedy when they refuse to engage in wrongdoing or report illegal practices, many state legislatures have passed some form of whistleblowing statute to create a statutory tort regime. Currently, twenty states have enacted whistleblowing statutes with general application to private employers. Of these twenty states, most also recognize a common law public policy exception to the common law at-will employment rule. Only Florida, Maine, New York, and Rhode Island do not. In addition, the status of the public policy exception in Arizona since the passage of the Arizona Employee Protection Act is unclear.

Of the thirty-one jurisdictions (thirty states plus the District of Columbia) which do not provide comprehensive statutory protection for whistleblowers, twenty-six states recognize some form of judicially created public policy exception. However, the extent of the protection afforded under the public policy exception is often limited and is

97. Id. Federal anti-discrimination laws, which contain a burden-shifting scheme that requires employers to articulate a legitimate, non-discriminatory reason for an adverse employment action, can be seen as an encroachment into the at-will regime, as can provisions specified in employee handbooks, which, when those provisions are held enforceable, may create additional rights for employees. See McDonnell Douglas Corp. v. Green, 411 U.S. 792, 802 (1973).


99. Id. I consider the following states to have whistleblower protections of general application to public and private employees: Arizona, California, Connecticut, Florida, Hawaii, Illinois, Louisiana, Maine, Michigan, Minnesota, Montana, New Hampshire, New Jersey, New York, North Dakota, Ohio, Oregon, Pennsylvania, Rhode Island, and Tennessee. Id. States could be added to or subtracted from this list, depending on how one defines a “general” whistleblower statute. For example, the New York statute only applies to private employees if an employee “discloses, or threatens to disclose to a supervisor or to a public body an activity, policy or practice of the employer that is in violation of law, rule or regulation which violation creates and presents a substantial and specific danger to the public health or safety . . . .” N.Y. LAB. LAW § 740(2)(a) (McKinney 2002); see also Callahan & Dworkin, supra note 80, at 107–08 (containing earlier fifty-state survey).


certainly not uniform across these jurisdictions. Some states require that the public policy underlying a retaliatory discharge claim be “well-established,” “substantial and widely accepted,” “well-recognized and clear,” “fundamental and well-defined,” or “strong and compelling.” These additional requirements lead to differences among states.

The fundamental problem in all public policy cases lies in determining exactly what constitutes a specific, well-established, clear, and compelling public policy, and courts have taken various approaches to resolve this question. Sources of “public policy” may include only state statutes, state statutes and the state constitution, “legislative, administrative or judicial authority,” or the “customs and conventions of the people” along with state statutes, the constitution, and judicial decisions. Some states broadly categorize the type of conduct protected by the public policy exception by defining it as reporting any illegal act.

Just as the state public policy decisions are uneven, so are the state whistleblowing statutes. State whistleblowing statutes consist of a patchwork of provisions. While some provide fairly broad protection, others limit protection to a specific industry or area.

107. Grzyb v. Evans, 700 S.W.2d 399, 401 (Ky. 1985).
111. See Grzyb, 700 S.W.2d at 401.
116. See Conn. Gen. Stat. § 31-51m (2003) (covering an employee of “a person engaged in business who has employees” who discloses violations of law); Ohio Rev. Code Ann. §§ 4113.51–.53 (West 2001 & Supp. 2003) (covering an employee of any employer with one or more employees who reports violations of law when “the employee reasonably believes that the violation either is a criminal offense that is likely to cause an imminent risk of physical harm to persons or a hazard to public health or safety or is a felony”).
Employee Whistleblowing After Sarbanes–Oxley

State whistleblowing statutes also differ in the type of disclosure they protect,118 the manner of disclosure they require,119 and the remedies they provide.120 Judicial interpretation has restricted the scope of some statutes.121 However, in other instances, it has broadened a statute’s application.122

Plaintiffs bringing actions for retaliatory discharge in various states must also meet varying burdens of proof. Depending on the language of the specific state statute, the whistleblower may have to prove that the reported wrongdoing actually occurred.123 Other jurisdictions use a less strict “reasonable belief standard,” requiring the whistleblower to prove that he or she had a good reason to believe the wrongdoing had occurred and thus made the report in good faith.124 Several courts have said that “bad faith” reports should not receive protection.125

§ 197.285 (West 2004) (covering hospital and ambulatory surgical center employees).


120. See ALASKA STAT. § 18.60.089 (Michie 2002) (filing complaint with commissioner); HAW. REV. STAT. § 388-10 (2004) (fining and/or imprisoning employer or agent); W. VA. CODE ANN. § 6C-1-4 (Michie 2003) (establishing civil action).

121. See, e.g., McDonald v. Campbell, 821 P.2d 139, 144–45 (Ariz. 1991) (holding that application of ARIZ. REV. STAT. §§ 38-531 to 38-534 to employees of the Arizona State Supreme Court is unconstitutional).


123. See, e.g., N.Y. LAB. LAW § 740 (McKinney 2002) (protecting private employee who “discloses, or threatens to disclose to a supervisor or to a public body an activity, policy, or practice of the employer that is in violation of law, rule, or regulation which violation creates and presents a substantial and specific danger to the public health or safety . . .”); Bordell v. Gen. Elec. Co., 667 N.E.2d 922, 923 (N.Y. 1996) (holding that under section 740, proof of an actual violation is required).

124. See, e.g., OHIO REV. CODE ANN. § 4113.52 (West 2001 & Supp. 2003) (protecting employee who reports a violation of law when “the employee reasonably believes that the violation either is a criminal offense that is likely to cause an imminent risk of physical harm to persons or a hazard to public health or safety or is a felony . . .”).

New York provides perhaps the best example of the shortcomings in whistleblowing law before the passage of Sarbanes–Oxley. New York’s whistleblower statute protects only workers who report a problem that would endanger the health or safety of the public. Specifically, the statute prohibits retaliatory action against an employee who “discloses, or threatens to disclose to a supervisor or to a public body an activity, policy or practice of the employer that is in violation of law, rule or regulation which violation creates and presents a substantial and specific danger to the public health or safety . . . .”

Even though financial fraud can result in massive layoffs if the company becomes bankrupt and can lead to individual penury or a complete loss of retirement savings, New York courts do not view fraud as a danger to the public welfare within the definition of the whistleblower statute. For example, in *Leibowitz v. Bank Leumi Trust Co. of New York*, the court held that financial fraud was not a danger to health and safety. Thus, the court did not afford the protection of the statute to the employee who had reported fraudulent loan practices. Further, in *Murphy v. American Home Products Corp.*, the court declined to create a public policy exception where the plaintiff had reported “$50 million in illegal account manipulations of secret pension reserves which improperly inflated the company’s growth in

| 126. N.Y. LAB. LAW § 740(2)(a) (McKinney 2002). |
| 127. *Id.* |
| 129. *Id.* at 520. |
| 130. *Id.* at 516–18. |
income and allowed high-ranking officers to reap unwarranted bonuses from a management incentive plan . . ."132

Although fraud is not a “danger” in the traditional sense of safety problems with a nuclear reactor or toxic chemicals discharged onto farmland, white collar crimes such as fraud may create serious economic damage. For example, many Enron employees who invested heavily in Enron stock have lost their savings for retirement.133 Although it is perhaps not an immediate “danger,” employees who witness fraud should have some ability to stop the criminal activity without worrying about being fired.134

As the foregoing section has shown, state whistleblower law is murky, piecemeal, disorganized, and varies from jurisdiction to jurisdiction. The stories of Watkins and Cooper demonstrate the differences between various state approaches and the vagaries and inconsistencies of those approaches.

2. Federal Whistleblower Statutes

The pre-Sarbanes–Oxley federal approach to whistleblower protection is similarly piecemeal. If a private-sector employee reports a violation of a federal statute, either internally or to a law enforcement agency, that employee may receive protection.135 Whether the whistleblower receives any legal protection depends on whether the federal statute violated by the employer contains an anti-retaliation provision.136 Conceptually, the whistleblower provisions that accompany each statute can be seen as a way of strengthening the enforcement of that statute.

One beneficial component of federal whistleblower protection is that government employees are generally covered by civil service

132. Id. at 87, 89.


134. Although Sarbanes–Oxley now covers reports of fraud, the Act’s civil provisions only apply to publicly traded companies. Sarbanes–Oxley Act of 2002, Pub. L. No. 107-204, § 806(a), 116 Stat. 745, 802–04 (codified at 18 U.S.C.A. § 1514A(b)(1)(B) (West Supp. 2003)). Therefore, an employee who reports fraud in New York may or may not have a civil claim, depending wholly on whether the employee works at a privately held company, a non-profit, or a publicly traded company. Although the public company requirement by necessity would probably encompass some of the state’s larger employers, the law still leaves a major gap.


136. Id.
For example, in the recent scandal over the abuse of Iraqi prisoners at Abu Ghraib prison, it was a whistleblower, Joseph Darby, who brought the problems to light. By slipping photographs of the abuse under a superior’s door, Darby brought necessary attention to the problem. His actions prompted an investigation, popular outcry, and ongoing court-martials of those involved. As a member of the military reporting torture and sexual abuse, Darby would be protected by a federal statute that specifically addresses such matters. Indeed, his conduct has earned him the respect of senior military officials. However, Darby’s family remains concerned about possible impacts on his career and personal life, as well as reaction from the soldiers who have been accused of the abuses. Despite these difficulties, government employees are generally much more well-protected than private employees.

Although Appendix B catalogues the federal statutes that include whistleblower protections, many others contain no such protection for whistleblowers. For example, if a private sector employee, such as a private contractor at a nuclear power plant, reports a violation of nuclear safety regulations, that individual would be covered by federal law because the regulations include an anti-retaliation provision. If, on the other hand, the statute is not one of those listed in Appendix B, it does not contain an anti-retaliation provision, and the job of a whistleblower...
reporting a violation of that statute to enforcement officials is not necessarily secure. For example, Title IX, which promotes equality in women’s and men’s educational opportunities, does not contain an anti-retaliation provision. Recently, a basketball coach alleged that after he reported unequal treatment for female athletes, he was fired. The United States Court of Appeals for the Eleventh Circuit rejected the claim, and the U.S. Supreme Court has granted certiorari.

In effect, the federal statutory scheme results in a haphazard enforcement structure. Whistleblower advocates continue to lobby Congress to pass a statute that provides a general federal cause of action for private-sector employees who are fired after blowing the whistle on a violation of federal law. Such a blanket anti-retaliation provision would make whistleblower protection the default rule for federal legislation.

3. Sociological Approaches to Whistleblowing

From this view of whistleblowing and employment law before the passage of the Sarbanes–Oxley, this Article briefly turns to a discussion of the sociology and psychology surrounding whistleblowers. It is important to understand that not only has the law been generally unsympathetic to whistleblowers, but so have co-workers and others outside the organization who do not support a decision to report wrongdoing. Watkins has said whistleblowing at Enron was “not an easy road to take” partly because the company planned to “just stick [her] in a corner and treat [her] like a pariah and sort of force [her] out. They just imploded too fast to have that plan work.” In light of this

150. Id.
151. One such advocacy group, the National Whistleblower Center, suggests just such a broad federal protection for whistleblowers on its website, which includes the text of a model act. See The National Whistleblower Protection Act, at http://www.whistleblowers.org/html/model_whistleblower_law.html (last visited Oct. 7, 2004). It is worth noting that such a model act would also be beneficial on a state level, given that the same issues of piecemeal coverage arise on the state level as well.
152. For an example of colleagues who did not support a whistleblower, see the case of Grace Pierce, a doctor on a research team who raised concerns about the use of saccharine in drugs for children. Pierce v. Ortho Pharm. Corp., 417 A.2d 505, 507 (N.J. 1980); see MYRON PERETZ GLAZER & PENINA MIGDAL GLAZER, THE WHISTLEBLOWERS 93–95 (1989).
extreme societal disapproval, any law that hopes to be effective in promoting whistleblowing must be a strong one.

One paradigm regarding whistleblowers casts the reporting employee as a type of scoundrel and inevitably focuses on the employee’s disloyalty to the organization. In this archetype, the whistleblower is a disgruntled employee with an axe to grind with his or her former employer. The whistleblower may well have been fired for incompetence, or let go in a layoff. In either event, the whistleblower feels the need to “get back” at his or her former employer, by implicating a co-worker or supervisor against whom he or she held a grudge.

On the other hand, whistleblowers are also portrayed as lone voices of reason, morality, and truth who speak out to protect the public from harm. This paradigm usually pits the conscience of one individual against the power and resources of a large organization. One need only watch the movies *Silkwood* or *The Insider* to see the myth of the heroic whistleblower writ large. In part, the myth of the heroic whistleblower represents American individualism. The individual worker, refusing to give up his or her morals and identity, instead stands up for what he or she believes is right in the face of overwhelming power and pressure to conform.

---

154. See Sherron Watkins, Address at Samford University (Feb. 19, 2004).
155. See Jones v. Enterprise Rent A Car, 187 F. Supp. 2d 670, 673 (S.D. Tex. 2002) (detailing plaintiff’s allegations that she was forced to perform various types of consumer fraud in order to keep her job, but later became the target of company investigation). A high profile example of whistleblower-as-wrongdoer was the case of Mark Whitacre, an executive at Archer Daniels Midland Co. who blew the whistle on anti-trust violations at the company. See generally KURT EICHENWALD, THE INFORMANT: A TRUE STORY (2001). Whitacre was later found to have embezzled over $9 million. Id.
Employee Whistleblowing After Sarbanes–Oxley

The scholarship on whistleblowers reveals quite a different story from either the “hero” or “dark side” scenarios—one in which whistleblowers are isolated by co-workers, relegated to dead-end positions, and affected in their personal lives by “spillover” from workplace stress associated with whistleblowing. According to a study of eighty-four whistleblowers conducted in the early 1990s, “82% experienced harassment after blowing the whistle, 60% were fired, 17% lost their homes, and 10% admitted to attempted suicide.” Due to the extreme stress, many whistleblowers develop serious mental illness, such as depression, which can lead to other problems, such as alcohol or drug abuse. The harsh truth about whistleblowing and the high price that whistleblowers pay is depicted in literature as well, most notably in Henrik Ibsen’s perceptive drama, An Enemy of the People.

In Bowen v. Parking Authority of the City of Camden, the whistleblowers confronted just such pressures. The two plaintiffs reported alleged corruption within the department, including the use of city property for personal gain, the distribution of materials that could be sexually harassing, and the inappropriate award of contracts based on the bidder’s personal “connections” and willingness to participate in a kickback scheme. While the two plaintiff-whistleblowers were in the process of cooperating with authorities and building a case, the stress mounted. At one point, one of the whistleblowers wrote a letter to the Equal Employment Opportunity Commission (EEOC) stating:

Over the last two years I have made diligent efforts to report and resolve numerous flagrant acts of discrimination/sexual harassment at our facilities . . . . At the present time I am functioning under intense pressure and threatened with termination . . . . Besides living in fear of losing my job I am

159. Culp, supra note 4, at 113.
160. Id.
161. Id.
162. See generally HENRIK IBSEN, An Enemy of the People, in THREE PLAYS OF HENRIK IBSEN (Elenor Marx-Aveling & William Archer trans., Heritage Press 1965) (1882). In the drama, the protagonist is confronted with a terrible choice: blow the whistle on the polluted water in his town and be fired from his job, evicted from his lodgings, and socially shunned, or keep silent, knowing all the while as a medical doctor that the sick and invalid who come to use the bath ostensibly for a cure are being made ill. Id.
164. Id. at *6.
165. Id. at **1–6.
166. Id. at *6.
also concerned for my personal safety as well as the safety of my co-workers . . . .

The second whistleblower also suffered from stress, anxiety, and depression. After being constructively discharged, the first whistleblower received a letter charging him with retaliation and harassment, apparently for making the reports. Tragically, the parking authority’s executive director, who had been deeply involved in the fraud and kickback schemes, shot and wounded the first whistleblower before turning the weapon on himself.

Despite negative reaction from supervisors and co-workers, it is often forgotten that the whistleblower is generally not being disloyal to the “true” goals of the organization. The organization’s stated goals do not (indeed, under corporate law principles, they cannot) involve illegal activity or fraud. Thus, although perhaps disloyal to management or an immediate supervisor who is performing illegal acts, whistleblowers are true to the stated ideals of their organization. For example, many WorldCom employees blamed Cynthia Cooper for her discovery and reporting of the WorldCom fraud, but they failed to realize that Cooper actually prevented further harm to the company and its employees by doing her job correctly. If the books can be corrected and the wrongdoers ousted from the organization, then the company may avoid further criminal sanctions and has a better chance to recover from the fraud.

Ultimately, employees receive little encouragement to blow the whistle. State and federal laws are inconsistent in their coverage, application, and enforcement. Additionally, the negative views of whistleblowing further discourage employees from reporting wrongdoing. With this sociological view in mind, this Article turns to an examination of the most salient provisions of Sarbanes-Oxley, which affect both securities law and employment law.

167. Id.
168. Id.
169. Id. at *8.
170. Id. at *14.
171. A corporation is empowered to act only to the extent that its business is lawful. Otherwise, the corporation is acting ultra vires. ROBERT CHARLES CLARK, CORPORATE LAW § 16.1, at 676 (1986) (citing MODEL BUS. CORP. ACT § 3.01 (1979)).
172. Lacayo & Ripley, supra note 2, at 33.
Employee Whistleblowing After Sarbanes–Oxley

II. SARBANES–OXLEY ACT CHANGES TO SECURITIES LAW AND EMPLOYMENT LAW

Depending on one’s view, the Sarbanes–Oxley Act either completely revised the field of securities regulation or merely made superficial changes to the status quo.174 Passing the Act in response to the corporate scandals that erupted at Enron,175 WorldCom,176 Global Crossing,177 Adelphia,178 and Tyco,179 Congress aimed to reduce accounting fraud, police insider transactions, and ensure the integrity of analyst research.180 In short, the purpose of the legislation is to increase transparency in financial markets, which allows investors to rely on the accuracy of financial information. As one commentator observed, “the primary policy of the federal securities laws involves the remediation of information asymmetries, that is, equalization of the information available to outside investors and insiders.”181 Fundamentally, securities law places a priority on disclosure because securities themselves do not have fixed or inherent value.182

Whether Sarbanes–Oxley actually advances these central goals of securities regulation is a matter of some debate. Obviously, the law will

---

174. As Joseph Grundfest, an SEC commissioner from 1985 to 1990, observed in a symposium on the collapse of Enron:

We are still too close to the Enron and WorldCom frauds, I believe, to draw any firm conclusion regarding the wisdom of the Congressional response as reflected in . . . Sarbanes–Oxley . . . . Some observers will complain that Sarbanes–Oxley has gone too far while others will protest that it has not gone far enough.


175. See supra Part I.A.1.

176. See supra Part I.A.2.


179. See infra notes 220–221 and accompanying text.

180. See supra note 8.


182. Seligman, supra note 181, at 450.
be perceived differently by corporate lawyers who prepare the periodic filings required by the SEC, plaintiff’s securities litigators, and securities defense attorneys. In part, the success of the law will depend on the SEC’s willingness and ability (through a proposed and much-needed increase in resources)\(^1\) to enforce the new laws. The Act’s success will also depend on the extent to which it is resisted and co-opted by industry and professional groups with major financial interests in the securities markets, such as auditors, corporate attorneys, investment bankers, and research analysts. Another possible barrier to the Act’s success is that corporate boards themselves are composed of rather insular groups—members have similar sets of connections and are often prone to groupthink. Will these groups actually listen to the employees who are blowing the whistle?\(^3\)

\(A. General Provisions\)

One of the major purposes of Sarbanes–Oxley is to promote the flow of accurate information to investors so that they can make informed decisions about how to allocate their resources.\(^5\) The whistleblowing provisions advance this purpose in that, if effective, they will reduce the amount of fraudulent financial information. The following section is a brief summary of the problems that gave rise to the Act’s provisions. In addition, it describes the measures that the Act and the regulations under the Act take to correct these problems, and also describes events subsequent to the passage of the Act.\(^6\)


184. In a way, the perspective of the whistleblower, who is guarding the underlying ideals of the organization, would be an ideal addition to a corporate board. How one would go about obtaining such a perspective, however, given the current system of corporate governance, is a difficult question.


Employee Whistleblowing After Sarbanes–Oxley

1. Research Analysts and Conflicts of Interest

Section 501 of the Act attempts to lessen conflicts of interest among research analysts. There has been considerable concern that research analysts working for investment banks were assigning positive “buy” ratings to the stock of corporations that used their employer’s investment banking services. Sometimes research analysts were given bonuses based on the amount of business the investment banking divisions received. As the analysts were ostensibly “independent,” this hidden conflict of interest tainted much of the research that had been released to the public. An investor could receive biased information without any indication that the source had an ulterior motive—that is, touting stocks underwritten by an affiliated investment bank.

The Act attempts to remedy these conflicts in two ways. First, the Act attempts to prevent investment banks from exercising control over research analysts. Section 501 prohibits investment banks from pre-approving research reports, supervising analysts, or compensating analysts in any way tied to their underwriting. Further, the investment banks may not retaliate against analysts for producing a report that is

188. See supra note 188.
189. See supra note 188.
190. See supra note 188.
191. See supra note 188.
193. Id.
neutral or negative.\textsuperscript{194}

The second way the Act seeks to prevent conflicts is through a regime of disclosure.\textsuperscript{195} Research reports must disclose whether that issuer is a client of the associated investment bank.\textsuperscript{196} In addition, analysts must disclose their own personal holdings in the issuer.\textsuperscript{197} Further, research reports must disclose the business the investment bank conducted with that company.\textsuperscript{198} Sarbanes–Oxley also requires investment banks to provide standardized criteria for giving a company a buy, sell, or hold rating.\textsuperscript{199} Finally, the Act prohibits a company from awarding bonuses to analysts based on the company’s level of investment banking business.\textsuperscript{200}

Since Congress passed the Act, there have been further developments to address conflicts of interest. On April 29, 2003, ten major investment banks announced that they would pay $1.4 billion into a fund to settle lawsuits accusing them of intentionally biasing their research.\textsuperscript{201} Of the total settlement, the parties agreed to place $432.5 million into a fund that would distribute research to investors independent of any brokerage firm.\textsuperscript{202} Additionally, many investors have brought securities arbitration actions on the basis of tainted research reports.\textsuperscript{203}

2. Increased Criminal Penalties for Securities Fraud

The Act augments existing white collar criminal statutes by creating two new criminal offenses, amending existing criminal statutes, and increasing the penalties for existing federal crimes, either by directly amending the statute or by requiring rulemaking to change the

\textsuperscript{194.} Id.
\textsuperscript{195.} Id.
\textsuperscript{196.} Id.
\textsuperscript{197.} Id.
\textsuperscript{198.} Id.
\textsuperscript{199.} Id.
\textsuperscript{200.} Id.
\textsuperscript{202.} Id.; see also SEC v. Bear, Stearns & Co., No. 03.CIV.2937 (WHP), 2003 WL 21517973, at *1 (S.D.N.Y. June 2, 2003) (describing settlement of cases and discussing continuing issues after the settlement).
\textsuperscript{203.} See, e.g., Susanne Craig & Ruth Simon, \textit{Morgan Stanley Is Ordered to Pay Client $100,000}, \textit{Wall St. J.}, June 13, 2003, at C7 (describing one such case).
Employee Whistleblowing After Sarbanes–Oxley

sentencing guidelines for those federal offenses. Two commentators have recently suggested that this approach adds little to the current federal laws regarding securities fraud. Another commentator, however, is more optimistic about the new criminal protections, stating that “the Act’s criminal provisions make significant strides toward piercing the veil of corporate silence.”

3. Standardization of Accounting Practices

Additionally, Sarbanes–Oxley attempts to standardize certain accounting practices. In part, this standardization is accomplished by the creation of the Public Company Accounting Oversight Board (PCAOB), an entity that regulates the private audit firms that currently monitor public company accounting. The PCAOB sets certain quality control, independence, and other ethics standards for outside auditors. Of course, as noted by one commentator,

[w]hether this body achieves anything useful will depend substantially on who is appointed. Judging by the work of comparable bodies attempting to oversee the accounting and auditing professions in past decades, this will not be an easy job to do well. Enormous lobbying pressure from those industries can be expected to dampen any major initiatives.

Indeed, as predicted, the Board became mired in controversy when


205. Michael A. Perino, Enron’s Legislative Aftermath: Some Reflections on the Deterrence Aspects of the Sarbanes–Oxley Act of 2002, 76 St. John’s L. Rev. 671, 698 (2002) (“[Congress] trumpeted [Sarbanes–Oxley’s] new crimes and new enhanced penalties as providing significant deterrence for securities fraud. In reality, these provisions are unlikely to have much real impact on deterrence. They criminalize little additional conduct and do little to enhance real penalties.”); Recent Legislation, Corporate Law—Congress Passes Corporate and Accounting Fraud Legislation, 116 HARV. L. REV. 728, 734 (2002) (arguing that only tangible effect will be number of charges a prosecutor can bring, perhaps inducing more guilty pleas).


208. Id.

209. Cunningham, supra note 18, at 944.
the SEC named William Webster as its chair. Although Webster had held numerous government positions, including that of CIA director, Webster had no connection or particular expertise with accounting and, in fact, had been a director of U.S. Technologies, a failed Internet firm that had been sued by shareholders for securities fraud. Ultimately, Webster resigned from the PCAOB before it held its first meeting. Critics of the Bush administration pointed to Webster’s lack of accounting expertise as evidence that the administration was far from serious about corporate governance reform.

4. Corporate Governance

In one of its more publicized provisions, the Sarbanes–Oxley Act mandates that chief executive officers and chief financial officers of publicly traded companies review, sign, and take responsibility for the periodic reports the companies file with the SEC. However, this is not in any real sense a departure from previous law. Company executives were already supposed to ensure that the corporation made no material misstatements in the periodic filings, and shareholders could sue executives for either nondisclosure or false disclosure of facts in the filings. As one commentator put it, “investors believed that the certification requirement was a meaningless exercise for those intent on committing financial fraud.” On the other hand, such a certification requirement would make it easier for prosecutors to prove the requisite state of mind needed to obtain a criminal conviction.

Sarbanes–Oxley has other implications for corporate governance. One

---

211. Id.
213. See Wilke, supra note 210, at A1.
216. See id.
217. Aronson, supra note 186, at 143. Another commentator succinctly summarized the ambivalence about these rules as follows: “These provisions look to prevent CEOs and CFOs from hiding behind the defense of ignorance . . . . The rules may be sensible, but few knowledgeable people really believed those senior executives, and none but the most sympathetic or gullible absolved them from responsibility.” Cunningham, supra note 18, at 955–56.
major change is that the Act forbids loans to insiders. 218 Before, insiders co-mingled personal funds with corporate funds, and many corporations gave insiders loans on extremely favorable terms. 219 Recent controversy about such co-mingling of funds has focused on L. Dennis Kozlowski, the former head of Tyco, Inc., who threw a lavish (indeed, decadent) party in Italy for his wife’s birthday, which prosecutors assert was charged to the company as part of a supposed board meeting that was also taking place in Italy. 220 In addition, the allegation that Kozlowski used Tyco’s assets to pay for a $6,000 shower curtain in a bathroom used only by the family’s maid has led Kozlowski to become not only the subject of many tasteless jokes, but also, more seriously, the subject of criminal prosecution. 221

The Act also contains items that have long been considered “best practices” for publicly-traded companies and are not novel in any way except that they are now mandatory. For example, issuers are required to disclose insider trades in a more timely fashion, 222 as well as to disclose other developments that could have an effect on the value of the company’s stock price. 223 The Act also strongly suggests that a “financial expert” be a member of the issuer’s audit committee. 224 Finally, the Act requires management to file a report with the SEC detailing its internal controls to ensure that fraud does not go unnoticed. 225

The Act also seeks to remedy the problem of auditor conflicts by prohibiting auditors from providing other services to their audit clients. 226 During the 1990s, one of the most lucrative sources of revenue

221. Colleen DeBaise & Mark Maremont, Jurors Examine Costs of Décor Chez Kozlowski, WALL ST. J., Dec. 16, 2003, at B1; Colleen DeBaise, Newest ‘Tyco Gone Wild’ Video Is Out, and Jurors See $6,000 Shower Curtain, WALL ST. J., Nov. 26, 2003, at C1 (describing, in addition to shower curtain, the outrageous prices of other home furnishings, including a mirror valued at $103,000 and a Persian rug valued at $191,250, which were allegedly furnished at Tyco shareholders’ expense).
223. Id. § 409 (codified at 15 U.S.C.A. § 78m (West Supp. 2003)).
225. Id. § 404 (codified at 15 U.S.C.A. § 7262 (West Supp. 2003)).
226. Id. § 201 (codified at 15 U.S.C.A. § 78j (West Supp. 2003)).
for accounting firms was selling not only audit services, but also other consulting and legal services to their auditing clients. The cross-selling of services, however, led to the very real threat of entanglement and conflicts of interest because an accounting firm’s auditing branch would be hesitant to criticize the consulting work that another branch of the same firm had performed. It was this type of interlock that led to the indictment and subsequent failure of accounting firm Arthur Andersen in connection with its role as consultant and auditor for Waste Management, Global Crossing, and Enron. Additionally, the Sarbanes–Oxley Act contains document destruction provisions, which make it a federal offense to destroy documents of the type that Arthur Andersen destroyed when it was under investigation.

5. Securities Fraud Litigation

Only a few years before the accounting scandals, Congress took action to make it more difficult for plaintiffs to bring securities class action lawsuits. The enforcement of our securities laws, like the enforcement of many regulatory schemes, is partly the responsibility of government agencies—in the case of securities regulation, the SEC—and is in part left to plaintiffs who take on the role of private attorneys general. In 1995, Congress passed the Private Securities Litigation and Reform Act (PSLRA) with the express purpose of limiting actions for securities fraud. The PSLRA “threw numerous substantive and procedural roadblocks” in front of plaintiffs seeking to bring securities fraud actions. Among other numerous requirements, it raised the

227. See Barbara Ley Toffler & Jennifer Reingold, Final Accounting 49–51 (2003). One former Arthur Andersen partner has described the process of gaining additional revenue from consulting as “turning a blind eye to accounting standards in order to earn the goodwill and trust of the client, and squeezing the consultants into meetings as often as possible in hopes of getting more overall business.” Id. at 49.

228. See id. at 50–51.


230. Sarbanes–Oxley Act § 802 (codified at 18 U.S.C.A. § 1519 (West Supp. 2003)). In response to this change in the law, companies should implement a standard document retention policy. Id.


232. See Bucy, supra note 145, at 31–54.

233. Private Securities Litigation Reform Act § 101(b).

Employee Whistleblowing After Sarbanes–Oxley

pleading standard: in order to survive a motion to dismiss, a complaint must allege facts giving rise to a strong inference that the defendants acted with scienter.\textsuperscript{235} Plaintiffs had to allege such facts without the benefit of discovery, which the PSLRA stayed until the complaint had survived a motion to dismiss.\textsuperscript{236}

The Sarbanes–Oxley changes to securities fraud litigation only marginally ease the burden on plaintiffs or mitigate the effects of the PSLRA. For actions that were commenced after July 30, 2002, § 804 extends the statute of limitations from one to two years after discovery of the fraud and from three to five years after the fraud occurred.\textsuperscript{237} Further, § 803 makes it easier for plaintiffs’ attorneys to collect restitution for securities fraud from the personal assets of corrupt executives.\textsuperscript{238} The Act does this by amending the bankruptcy code to prevent the discharge in bankruptcy of debts related to securities fraud.\textsuperscript{239} Finally, § 306 gives private litigants the right to bring a derivative lawsuit to recover profits that executives earn through insider trades made during pension fund blackout periods (i.e., if pension funds are frozen out and cannot trade in the stock, executives should not be allowed to trade in the stock at that time either).\textsuperscript{240}

B. Whistleblowing Provisions

In addition to the changes described above, Sarbanes–Oxley attempts to respond directly to the plight of employees who blow the whistle on corporate fraud. The following section details the changes that Sarbanes–Oxley makes to securities law and employment law. In particular, Sarbanes–Oxley provides federal statutory protection to whistleblowers who report fraud at publicly traded companies.\textsuperscript{241}

\begin{thebibliography}{99}
\bibitem{Bites} Bites Posing as Reform, 12 A.B.A. SEC. NEWS 2, 2 (Winter 2003).
\bibitem{ scienter} 15 U.S.C. § 78u-4(b)(2).
\bibitem{pleading} Id. § 78u-4(b)(3)(B). One consequence of the PSLRA is that pension funds and other institutional investors must take a leading role in prosecuting securities class action litigations. Arden Dale, Pensions Join Class-Action Suits at Faster Pace, Lending Clout, WALL ST. J., Jan. 14, 2004, available at 2004 WL-WSJ 56917066.
\bibitem{easier} See id. § 803 (codified at 11 U.S.C.A. § 523(a) (West Supp. 2003)).
\bibitem{protection} Id. § 806(a) (codified at 18 U.S.C.A. § 1514A (West Supp. 2003)).
\end{thebibliography}
provides criminal penalties for retaliation against whistleblowers, and requires publicly traded companies to institute procedures for handling internal complaints.

1. Section 806: Federalizing the Law of Whistleblowing

Under § 806 of Sarbanes–Oxley, whistleblowers who report instances of fraud internally or to governmental agencies are statutorily protected from retaliation if they work at publicly traded companies. Sarbanes–Oxley therefore changes the landscape of whistleblower law by federalizing a portion of the law that had been composed of a patchwork of federal statutes, state statutes, and common law exceptions to the at-will employment rule. Had it been in force before the scandals broke, § 806 would have been directly applicable to both Watkins’s and Cooper’s employment situations.

Section 806 expressly forbids a publicly traded company or “any officer, employee, contractor, subcontractor, or agent of such company” from discharging, demoting, suspending, threatening, harassing, or otherwise discriminating against an employee who engages in certain forms of protected whistleblowing activity. The Act defines protected activity in one of two ways. First, employees are protected when they “provide information, cause information to be provided, or otherwise assist in an investigation regarding any conduct which the employee reasonably believes constitutes a violation of . . . any provision of Federal law relating to fraud against shareholders . . . .” This type of activity is protected, however, only if the “information or assistance” is provided in an appropriate forum: to a “Federal regulatory or law enforcement agency,” to a “[m]ember of Congress or any committee of Congress,” or to “a person with supervisory authority over the employee (or such other person working for the employer who has the authority to investigate, discover, or terminate misconduct) . . . .” Alternatively, the Act protects employees when they:

file, cause to be filed, testify, participate in, or otherwise assist
in a proceeding filed or about to be filed (with any knowledge of the employer) relating to an alleged violation of . . . any rule or regulation of the Securities and Exchange Commission, or any provision of Federal law relating to fraud against shareholders.\textsuperscript{249}

In other words, the Act protects whistleblowers who make internal reports of violations, as long as those reports are made to a supervisor or another individual within the organization.\textsuperscript{250} Moreover, the supervisor must have enough institutional power to investigate the wrongdoing.\textsuperscript{251} The Act, therefore, does not protect an employee who confers with a peer, or discusses an accounting impropriety with a subordinate internally.

Externally, the law covers reports to government agencies, such as the SEC.\textsuperscript{252} It also protects employees who provide information that would assist a shareholder lawsuit for a securities fraud violation.\textsuperscript{253} Reporting a violation to the media, however, would not be protected. One would assume that, although a report to the media would effectively expose the relevant concern, such a report could result in an extreme amount of adverse publicity.\textsuperscript{254} One could view this as striking an economic balance between encouraging reports of wrongdoing and, at the same time, preventing mistaken, erroneous, or malicious reports from reaching the press, where substantial harm to a company’s reputation—not to mention stock price—would result.\textsuperscript{255}

The whistleblower’s required state of mind is described by the statute, but this standard will likely be interpreted by courts. The issue is whether the whistleblower’s motives must be pure or whether the whistleblower is still protected if he or she intended to implicate a co-worker or cover up his or her own wrongdoing. What if the complaints of fraud are completely false or without any basis in fact? Section 806 of the Sarbanes–Oxley Act establishes that, in order to obtain the protection

\textsuperscript{249}. Id.
\textsuperscript{250}. See id.
\textsuperscript{251}. Id.
\textsuperscript{252}. Id.
\textsuperscript{253}. Id.
\textsuperscript{254}. For a discussion of the issues surrounding whistleblowing to the media, see generally Elletta Sangrey Callahan & Terry Morehead Dworkin, \textit{Who Blows the Whistle to the Media and Why: Organizational Characteristics of Media Whistleblowers}, 32 AM. BUS. L.J. 151 (1994).
\textsuperscript{255}. The media lacks any formal institutional power to address the problem, unlike Congress, a regulatory agency, or a supervisor within organization.
of the statute, a whistleblower must have a “reasonable belief” that there has been a violation of the federal securities laws. If the state cases are any indication, this statutory language will likely become a subject for litigation.

If an employer violates § 806, the employee is entitled to “all relief necessary to make the employee whole.” This “make whole” relief includes “reinstatement with the same seniority status that the employee would have had, but for the discrimination,” “back pay, with interest,” and litigation costs, including attorney’s fees. There is, however, no provision for punitive damages.

The procedure for commencing a whistleblowing action is only briefly described in the Act. Instead, Sarbanes–Oxley adopts the procedures for whistleblowing actions set forth in another statute, the Wendell H. Ford Aviation Investment and Reform Act. The Sarbanes–Oxley Act gives oversight of § 806 to the Department of Labor (DOL), which has since delegated authority to the Secretary of the Occupational Safety and Health Administration (OSHA). Newly promulgated regulations further elaborate the procedures.

These regulations require the employee to send notification of the OSHA complaint to the employer. Within twenty days of the initial filing, the OSHA investigator must allow the alleged violator/employer the opportunity to respond. The investigation will occur only if the plaintiff establishes a prima facie case that the adverse employment action was the result of protected activity and not some other lawful personnel reason. Under a DOL decision, to make out a prima facie case the whistleblower must establish by a preponderance of the

259. Id.
260. See id.
261. See id.
265. Id. § 1980.104.
266. Id.
267. Id.
Employee Whistleblowing After Sarbanes–Oxley

evidence that (1) the whistleblower engaged in protected activity as defined by Sarbanes–Oxley, i.e., making a report as described in § 806; (2) the employer was aware of the protected activity; (3) the whistleblower suffered an adverse employment action; and (4) the whistleblower raised an inference that the protected activity was a contributing factor to the adverse employment action.\(^{268}\) If the plaintiff successfully establishes a prima facie case, the secretary will investigate and issue a determination in the form of a letter.\(^{269}\) If the secretary finds there is reasonable cause to believe there has been retaliation, then the secretary must award “make whole” relief.\(^{270}\)

Section 806 also contains a timeline for the administrative proceeding. If there has been no final decision within 180 days of the complaint, the plaintiff is entitled to de novo review in federal district court.\(^{271}\) One commentator has noted that this timeline, while ambitious, is also unrealistic and the end result may be that defendants will need to defend the lawsuit in front of two tribunals—once when they begin proceedings in an administrative forum, and once de novo in district court, where they will find themselves simply because of the timeline.\(^{272}\)

2. Section 1107: Criminal Penalties for Shooting the Messenger

Other portions of the Act also are concerned with the type of whistleblowing in which Watkins and Cooper engaged.\(^{273}\) Entitled


\(^{269}\) 29 C.F.R § 1980.105.

\(^{270}\) Sarbanes–Oxley Act § 806(a).

\(^{271}\) Sarbanes–Oxley Act § 806(a).


\(^{273}\) See Brickey, supra note 42, at 359–60 (describing impact of criminal provisions, including criminal whistleblowing provisions). Brickey concludes that:

Sarbanes–Oxley undoubtedly will not be the last word on corporate governance reform or punishing criminal fraud. There are no simple solutions to a culture of deceit fueled by greed, mismanagement, conflicts of interest, and failure of professional and regulatory oversight. But Sarbanes–Oxley is a constructive step in the right direction. It may fall short, but it sends an unmistakably clear signal that this should never happen again.

Id. at 381.
“Retaliation Against Informants,” § 1107 of the Act creates a new criminal offense of that same name. Section 1107 amends an existing federal statute that deals with violence against witnesses and makes it a crime to “knowingly, with the intent to retaliate [take] any action harmful to any person, including interference with the lawful employment or livelihood of any person, for providing to a law enforcement officer any truthful information relating to the commission . . . of any Federal offense . . . ” Before this amendment, the statute applied only to retaliation that caused death, bodily injury, or damage to property.

In addition to its application to whistleblowing about securities fraud and accounting violations, § 1107 also criminalizes retaliation against an employee who reports any federal offense to law enforcement. Further, the language of § 1107 is not limited to publicly traded companies. Indeed, its reach would seemingly encompass every employer. Section 1107 conceivably criminalizes many actions that have been wholly civil tort matters, such as retaliation against an employee who reports violations of Title VII to the EEOC.

One management-side labor and employment law attorney has referred to § 1107 as “[o]ne of the Act’s potentially most dangerous sections” because it imposes criminal penalties for run-of-the-mill whistleblowing cases. One feature that distinguishes § 1107 from the average whistleblower scenario is that criminal penalties attach only if there is an “intent to retaliate,” the information about the commission of the federal offense is “truthful information,” and it is provided to a “law enforcement officer.” Further, the effectiveness of this provision is wholly dependent on the discretion of the prosecutor, which is a separate

276. Id.; Victim and Witness Protection Act § 4(a), 18 U.S.C. § 1513 (2000). Perhaps the extension to employment can be rationalized by recognizing that employment relationship can on some level be considered a property right. For an example of a work that treats employment in this manner, see generally Joseph William Singer, The Reliance Interest in Property, 40 STAN. L. REV. 611 (1988).
278. See id.
Employee Whistleblowing After Sarbanes–Oxley

topic altogether.281

3. **Section 301: Handling Whistleblower Complaints**

Section 301 of the Act, which deals with audit committees, also affects whistleblowers.282 Section 301(4) requires each company’s audit committee to establish procedures for “the receipt, retention, and treatment of complaints received by the issuer regarding accounting, internal accounting controls, or auditing matters” and “the confidential, anonymous submission by employees of the issuer of concerns regarding questionable accounting or auditing matters.”283 This is not phrased in discretionary terms—every publicly traded company must have a system in place for receiving anonymous complaints.284

Of course, the “procedures” here are only briefly addressed within the Act, a flaw that will be discussed and analyzed in Part III of this Article. As with other sections of the Act, there is a question as to whether these briefly-mentioned “procedures” for anonymous complaints must be internal or external. The Act is silent on this point, and therefore many companies interpret this provision to mean that an internal system for reporting complaints would suffice.285

**III. THE EFFECTIVENESS OF SARBANES–OXLEY**

From the previous section it is clear that Sarbanes–Oxley certainly has, and will continue to have, an impact upon corporate governance and securities law. However, the magnitude of the impact is an open question, and one that the next section critically analyzes. How much of a change does the Act actually make, and just how much protection are whistleblowers afforded? Evaluating the changes to employment law

---

Courts recognize a prosecutor’s broad discretion to initiate and conduct criminal prosecutions, in part out of regard for the separation of powers doctrine and in part because ‘the decision to prosecute is particularly ill-suited to judicial review.’ So long as there is probable cause to believe that the accused has committed an offense, the decision to prosecute rests within the prosecutor’s discretion. A prosecutor has far-reaching authority to decide whether to investigate, grant immunity, or permit a plea bargain, and to determine whether to bring charges, what charges to bring, when to bring charges, and where to bring charges.

*Id.* (citations omitted).


283. *Id.*

284. *Id.*

285. *See infra note* 314 and accompanying text.
assists in determining the effectiveness of the Act. Both of these analytical and evaluative goals can be answered by an inquiry into the weaknesses of Sarbanes–Oxley.

After an examination of the Act’s weaknesses, this Article concludes that Sarbanes–Oxley should have provided more protection for whistleblowers. First, the Act lacks procedures for responding to a whistleblower’s report. Second, a whistleblowing case can apparently be sent to arbitration, a forum that this Article argues weakens the rights of employees. A detailed analysis of Sarbanes–Oxley supports the proposition that the Act, a product of political compromise, is far more limited than the rhetoric accompanying it would indicate.

Sarbanes–Oxley, in its provision of a civil scheme for recovery in § 806, is an area-specific statute dealing with fraud at publicly traded companies. Whistleblowing in other areas is still left with piecemeal state law coverage, and an employee is either protected or denied recovery based on the nature of the violation and the happenstance of the law in the relevant jurisdiction. Given the confusing status of state whistleblowing law, this Article concludes that a general federal whistleblower protection act would be better policy than the current patchwork of area-specific statutes. Under the reform proposed, an employee who reported any violation of state or federal law, rule, or regulation would receive federal civil protection from retaliation. Further, the law would make it clear that these disputes are not subject to mandatory pre-dispute arbitration. First, however, the following two sections of this Article examine the effectiveness of Sarbanes–Oxley’s whistleblower protections.

A. Sarbanes–Oxley Fails to Specify Procedures that Employers Must Follow when Addressing Employee Complaints

The Act fails to detail the procedures employers must follow when dealing with a whistleblower complaint. Part II.B.3 described § 301 of the Act, which requires that “[e]ach audit committee shall establish procedures” for “the receipt, retention, and treatment of complaints received by the issuer regarding accounting, internal accounting

287. For the text of a proposed act by the National Whistleblower Center, a whistleblower’s advocacy group, see The National Whistleblower Protection Act, supra note 151. It is worth noting that such a model act would be also be beneficial on a state level to provide employees uniform protections.
controls, or auditing matters” and “the confidential, anonymous submission by employees of the issuer of concerns regarding questionable accounting or auditing matters.”288 However, § 301 fails to specify what types of procedures are adequate.

Although the law requires a channel for employees to send anonymous complaints, and the audit committee is supposed to create procedures for the “retention” and “treatment” of complaints, the Act does not describe what the “treatment” of such complaints entails.289 Thus, like the proverbial tree in the forest, whistleblowers can report problems under the Act, but there is no guarantee that anyone—on the audit committee or otherwise—will necessarily hear them. This provision creates a “black hole” of sorts, in which anonymous complaints flow in, but there is no description of what to do with the complaints once they arrive.

In *Faragher v. City of Boca Raton*,290 the U.S. Supreme Court provided guidance on how best to deal with complaints stemming from instances of sexual harassment.291 The decision gives employers an affirmative defense to Title VII liability if they take prompt corrective action to stop harassment from occurring once it has been reported.292 It also shields employers from liability if they have in place policies or procedures that should prevent such harassment from occurring in the first place.293 No such analogue currently exists for Sarbanes–Oxley whistleblowing, but one should.

Presumably, in requiring the “treatment” of the anonymous complaints, § 301 implies that the employer must take at least some sort of action once it receives a complaint.294 However, because the Act does not set out specific requirements,295 conceivably the complaints could be “treated” by marking them with a large red check mark, or moving

288. Sarbanes–Oxley Act § 301.
289. Id.
291. Id. at 786–92.
293. *Faragher*, 524 U.S. at 807. Such measures might include training employees or providing human resources staff who know how to address such complaints.
295. See id.
electronic mail into a reviewed folder—the Act does not require the employer to respond to the complaint in any way. This is not because it is inherently impossible to respond. One can imagine instances where even if a complaint is sent anonymously, it is possible to let that employee know that the feedback has been received and is being responded to, perhaps by responding to the anonymous e-mail account or posting an announcement on an electronic bulletin board.

The investigation into mutual fund timing at Putnam Investments provides one example of the black hole: a whistleblower tried to report financial fraud but was utterly ignored even after Sarbanes–Oxley. 296 The whistleblower, Peter Scannell, reported that institutional investors were engaging in market timing, which is defined as: “[buying] shares of international funds late in the day when the U.S. market is rallying, locking in the daily fund price set at 4 p.m. Because foreign markets regularly follow the U.S., foreign shares in the fund will usually rise the next day, producing a higher fund price, enabling the trader to reap a profit by quickly selling thereafter.” 297 Although this procedure was technically legal at the time the scandal broke, it had the effect of harming long-term mutual fund shareholders and violated internal policies at Putnam. 298 Acknowledging that such trading hurts returns for shareholders, many funds had falsely stated in their prospectus or other materials sent to shareholders that they did not engage in timing. 299 Many of the calls for increased action in the regulation of mutual funds focus on these types of misstatements in fund disclosures to the public. 300

Scannell, an employee at Putnam’s call center in the suburbs of Boston, remarked on the fact that members of Boilermakers Union, Local 5, had made thousands of trades in international funds in violation

297. See Hechinger, supra note 296, at C13.
298. Id. Since the scandal broke, the SEC has begun considering options for new regulations that would make the practice of timing illegal. These proposed rules are already drawing controversy because they involve naming an employee of the mutual fund management company as a compliance officer. Karen Damato & Michael Schroeder, Can Funds Police Themselves?, WALL ST. J., Dec. 4, 2003, at C1.
299. See Hechinger, supra note 296, at C1.
300. See Damato & Lauricella, supra note 296, at C15.
Employee Whistleblowing After Sarbanes–Oxley

of Putnam’s internal policies.\textsuperscript{301} Upon bringing this to the attention of his two immediate supervisors, Scannell allegedly was told “[t]hat’s Putnam senior management’s concern.”\textsuperscript{302} After being ignored internally and then attacked by a man wielding a brick (according to a police report, the assailant was wearing a Boilermakers T-shirt), Scannell eventually turned to the SEC (who initially ignored him) and then to Massachusetts regulators.\textsuperscript{303} Regulators have now brought a civil fraud claim against Putnam in state court.\textsuperscript{304}

Perhaps the only tangible effect of § 301 is that it may well provide concrete, written evidence, either for government investigators or for plaintiff’s attorneys prosecuting shareholder derivative lawsuits, of who knew about a fraud and when they decided to ignore or tacitly assent to it. A common problem that arises when prosecuting shareholder lawsuits is that defendants “pass the buck” among themselves. One defendant will point his or her finger at another corporate defendant and argue that someone else accounted for revenues improperly or masterminded the fraud.\textsuperscript{305} Similar to this argument is the “rogue trader” defense, in which a high-ranking executive will argue that he or she was ignorant of a fraudulent practice because it was the fault of an out-of-control “rogue” subordinate.\textsuperscript{306} Of course, the typical response\textsuperscript{307} to this argument is that

\begin{thebibliography}{9}
  \bibitem{note301} Hechinger, \textit{supra} note 296, at C13.
  \bibitem{note302} \textit{Id.} Although in this instance the Boilermaker’s Union was responsible for the trades, Putnam apparently admitted that “six of its managers made $700,000 in such trading.” See John Hechinger & David Armstrong, \textit{Mutual Funds Vow to Fix Their Clocks: Putnam Says it Was Subpoenaed for Papers Tied to Fund Trading; CEO Subject of Internal Scrutiny}, \textit{WALL ST. J.}, Oct. 28, 2003, at C12.
  \bibitem{note303} Hechinger & Armstrong, \textit{supra} note 302, at C12.
  \bibitem{note304} \textit{Id.}
  \bibitem{note305} In \textit{In re Symbol Technologies}, many of the defendants began implicating each other in the fraud. Memorandum of Law in Support of Defendant Tomo Razmilovic’s Motion to Dismiss the Consolidated Amended Class Action Complaint at 9, 10, \textit{In re Symbol Technologies}, Inc. (E.D.N.Y 2003) (No. 02-CIV-1383 (LDW)) (settled without opinion); Defendant Frank Borghese’s Motion to Dismiss at 4, 12, 13, \textit{In re Symbol Technologies}, Inc. (E.D.N.Y. 2003) (No. 02-CIV-1383 (LDW)) (settled without opinion). The case eventually settled for $139 million, and the former CEO became a fugitive after his indictment. Michael Weissenstein, \textit{Indicted Executive Declared Fugitive}, \textit{ASSOCIATED PRESS NEWSWIRE}, June 9, 2004, available at 2004 WL 82210148.
  \bibitem{note306} In \textit{In re Reliant Securities}, the defendants argued that so-called “round trip” energy trades were the products of subordinates rather than management. Reliant Defendants’ Motion to Dismiss Plaintiffs’ Claims Under Sections 10(B) and 20(A) of the Securities Exchange Act of 1934, Section 12(A) of the Securities Act of 1933, and the Texas Securities Act, at 12, 20, \textit{In re Reliant Securities} (S.D. Tex. 2003) (No. H-02-1810).
  \bibitem{note307} Securities fraud lawsuits seem to contain the same archetypal arguments. In response to a plaintiff’s complaint alleging fraud, defendants will almost always counter with an argument that
\end{thebibliography}
a “rogue” subordinate may only engage in such behavior under inadequate supervision and a lack of proper internal controls. 308

It is difficult to discern public companies’ response to § 301. Before the passage of the Sarbanes–Oxley Act, some companies had, on their own, set up internal “ethics hotlines,” which make sense as a best practice from a purely economic perspective. 309 At the time Enron was formulating and implementing its byzantine accounting structures, it had in place a phone hotline that employees could call to report fraud anonymously. 310 Clearly this hotline did not prevent the massive fraud. The problem, as Sherron Watkins frames it, is that managers were continuously pushing the edge between acceptable and non-acceptable accounting practices and that, because of the company culture, no one wanted to challenge the hierarchy. 311 Watkins posits that no one called the hotline because the border between what was acceptable and what was illegal had been blurred, and the employees were afraid to speak up about the situation for fear that they would lose their jobs. 312 Watkins herself initially turned to an anonymous suggestion box to air her concerns, but then went directly to Ken Lay to discuss them. 313

To comply with § 301, many publicly traded companies have hired companies that specialize in setting up either anonymous hotlines or electronic mail systems that employees can use to report fraud. 314
Employee Whistleblowing After Sarbanes–Oxley

Common sense, as well as the results of one worker survey, indicates that people perceive anonymous hotlines as more “trustworthy” when they are maintained and monitored by a third-party,” and therefore those hotlines receive more calls than internal hotlines. Given the disapproval and reprisals that whistleblowers face, it is not surprising that some workers are reluctant to give up their anonymity. Thus, while § 301 improves the pre-existing regime, which did not require any action whatsoever of the employers, it is still not likely to be effective and certainly is not a revolutionary change.

B. Mandatory Arbitration and Whistleblowing Claims

The effectiveness of any statute depends on the remedies available to successful plaintiffs. Therefore, it is necessary to analyze the remedies whistleblowers are entitled to in the event retaliation occurs. While § 806 of the Sarbanes–Oxley Act authorizes “make whole” relief to employees who have been fired because of protected whistleblowing activity, many of these cases may never reach an administrative tribunal or a federal court. Rather, many Sarbanes–Oxley whistleblowing cases, like many employment discrimination cases, will probably be sent to mandatory arbitration. My earlier work describes many of the problems associated with mandatory arbitration: perceived employer bias; fewer options available to the employee regarding where

Services, which runs a more traditional phone-based reporting system); Charles Lunan, Law Prompts Business Boom: Accountants, Lawyers and Ethics Consultants Rush to Help Companies Comply with the Sarbanes–Oxley Reform Act, CHARLOTTE OBSERVER, May 25, 2003, at 1E (reporting growth of Pinkerton Compliance Services, including an increase from 780 public-company clients to 900 in the last year); see also Whistleblower Hotline, at http://www.shareholder.com/home/Solutions/Whistleblower.cfm (last visited Oct. 7, 2004) (advertising secure hotline system free of “human intervention”).

315. Ernst & Young, supra note 5 (reporting results of survey that concluded that while only thirty percent of workers would report fraud to management by means of an anonymous telephone call, that percentage rose by an additional twenty-seven percent if there was a telephone hotline administered by a third-party).
316. See supra Part I.B.3.
to bring suit; and limitations on discovery.\textsuperscript{319} All of these factors decrease the settlement value of a plaintiff’s case and, as a result, undermine the federal protections granted to employees.\textsuperscript{320} The Sarbanes–Oxley Act fails to address this issue, and that failure is yet another weakness of the Act.

1. \textit{Arbitration and the U.S. Supreme Court: From Skeptic to Supporter}

Before the early 1990s, courts did not enforce pre-dispute mandatory arbitration contracts for statutory employment claims.\textsuperscript{321} However, in recent years, there has been a complete doctrinal reversal, and courts now view mandatory arbitration contracts favorably.\textsuperscript{322} The U.S. Supreme Court cases decided during the last twenty years are the most critical to understanding the current state of the law on mandatory arbitration contracts. In the 1974 decision \textit{Alexander v. Gardner-Denver Co.},\textsuperscript{323} the Court refused to send a Title VII claim to mandatory arbitration.\textsuperscript{324} However, the Court later endorsed mandatory arbitration of statutory claims in \textit{Gilmer v. Interstate/Johnson Lane Corp.}\textsuperscript{325} and \textit{Circuit City Stores, Inc. v. Adams}.\textsuperscript{326}

In \textit{Gardner-Denver}, the Supreme Court unanimously held that a plaintiff suing under Title VII for race discrimination retained an independent right to be heard in federal court, despite his or her participation in an arbitration scheme established under a collective bargaining agreement.\textsuperscript{327} The Court reasoned that restricting plaintiffs to arbitration would interfere with both the EEOC and the individual litigant’s role as a private attorney general in enforcing Title VII.\textsuperscript{328} The pervasive line of reasoning in the opinion, however, is that the federal courts have plenary power over, and are the most competent to hear,

\begin{itemize}
\item \textsuperscript{319} Cherry, \textit{supra} note 318, at 276–86.
\item \textsuperscript{320} \textit{Id.}
\item \textsuperscript{321} \textit{See} \textit{Alexander v. Gardner-Denver}, 415 U.S. 36, 48 (1974).
\item \textsuperscript{323} 415 U.S. 36 (1974).
\item \textsuperscript{324} \textit{Id. at} 59–60.
\item \textsuperscript{325} 500 U.S. 20 (1991).
\item \textsuperscript{326} 532 U.S. 105 (2001).
\item \textsuperscript{327} \textit{Gardner-Denver}, 415 U.S. at 59–60.
\item \textsuperscript{328} \textit{Id. at} 44–45 (“Congress gave private individuals a significant role in the enforcement process of Title VII . . . . [T]he private litigant not only redresses his own injury but also vindicates the important congressional policy against discriminatory employment practices.”).
\end{itemize}
Employee Whistleblowing After Sarbanes–Oxley

Title VII claims.\textsuperscript{329} The Court stated that “final responsibility for enforcement of Title VII is vested with federal courts.”\textsuperscript{330} The Court compared Title VII litigation and arbitration and ultimately held exclusive arbitration proceedings inappropriate for resolving Title VII claims.\textsuperscript{331} The Court cited the limited discovery, the inapplicability of rules of evidence, and the overall informality of an arbitration proceeding as defects that prevented arbitration from being an exclusive remedy for a Title VII plaintiff.\textsuperscript{332} In addition, the Court also pointed to the arbitrator’s lack of “general authority to invoke public laws that conflict with the bargain between the parties”\textsuperscript{333} and possibly limited expertise with employment discrimination lawsuits.\textsuperscript{334} Taken as a whole, the \textit{Gardner-Denver} decision evinced a strong hostility to the exclusive arbitration of Title VII claims.

The \textit{Gilmer} decision, in 1991, involved a claim under the Age Discrimination in Employment Act (ADEA).\textsuperscript{335} The plaintiff, a manager in the securities industry, had signed a pre-dispute mandatory arbitration clause as required by the New York Stock Exchange.\textsuperscript{336} The Court held that the arbitration clause was enforceable and could constitute an exclusive remedy.\textsuperscript{337} The Court, however, did not overrule \textit{Gardner-Denver}, but chose instead to distinguish it by stressing that \textit{Gardner-Denver} involved a collective bargaining agreement.\textsuperscript{338} According to the \textit{Gilmer} Court, the operative distinction in \textit{Gardner-Denver} was that the union had traded away the individual plaintiff’s rights under Title VII.\textsuperscript{339} In \textit{Gilmer}, no such collective/individual tensions were present because it

\textsuperscript{329.} Id. at 45.
\textsuperscript{330.} Id. at 44.
\textsuperscript{331.} Id. at 56–60.
\textsuperscript{332.} Id. at 57–58.
\textsuperscript{333.} Id. at 53.
\textsuperscript{334.} Id. at 57 (“[T]he specialized competence of arbitrators pertains primarily to the law of the shop, not the law of the land.”).
\textsuperscript{335.} Gilmer v. Interstate/Johnson Lane Corp., 500 U.S. 20, 23 (1991). For examples of pre-\textit{Gilmer} decisions where lower courts generally assumed that Title VII claims, whether in the collective bargaining or contractual context, were non-arbitrable under \textit{Gardner-Denver}, see Utley v. Goldman Sachs & Co., 883 F.2d 184, 185 (1st Cir. 1989); Swenson v. Mgmt. Recruiters Int’l, Inc., 858 F.2d 1304, 1305–07 (8th Cir. 1988).
\textsuperscript{336.} Gilmer, 500 U.S. at 23.
\textsuperscript{337.} Id.
\textsuperscript{338.} Id. at 35.
\textsuperscript{339.} See id.
involved merely a contract with an individual employee, not a union. This result seems somewhat paradoxical, as the plaintiff in *Gilmer* probably had less bargaining power due to the absence of a union.

Then in *Circuit City*, a six-to-three decision, the Supreme Court spoke definitively about the validity of mandatory arbitration in the employment context. After analyzing the text of the Federal Arbitration Act (FAA), the Court concluded that the exception contained within the FAA for certain contracts of employment applied only to transportation workers. In making this determination, the court stated that “arbitration agreements can be enforced under the FAA without contravening the policies of congressional enactments giving employees specific protection against discrimination prohibited by federal law . . .”

This pro-arbitration decision was tempered by the Court’s opinion in *Equal Employment Opportunity Commission v. Waffle House, Inc.* There, the Court held that even though the plaintiff had signed a mandatory arbitration contract, the contract did not preclude the EEOC from pursuing an enforcement action for victim-specific relief. One major criticism of arbitration has been that it tracks employees raising civil rights violations into individualized arbitrations, as opposed to class actions. This sort of tracking apparently precluded any hope of correcting systemic civil rights violations in the workplace, and many commentators had previously criticized arbitration on that basis alone.

In *Waffle House*, the Supreme Court apparently recognized the limitations of private contracts in the face of a statutory and regulatory scheme to end disability discrimination.

However, on the whole, the overwhelming impression that one derives from the *Circuit City* and *Waffle House* decisions is that courts have shifted from viewing mandatory pre-dispute arbitration contracts with suspicion to distinctly favoring them. In part this can be explained

---

340. *Id.* at 23.
343. *Circuit City*, 532 U.S. at 119.
344. *Id.* at 123.
346. *Id.* at 297.
348. See, e.g., *id.*
Employee Whistleblowing After Sarbanes–Oxley

by the fact that enforcing mandatory arbitration contracts frees the federal courts from having to deal with a considerable volume of employment litigation.\textsuperscript{349} Arbitration of consumer claims and employment claims, among others, has become more widespread in response to increased enforcement of mandatory arbitration contracts.\textsuperscript{350} In light of this doctrinal background, a statute must clearly preclude mandatory arbitration, or the court will assume that the statutory claim is one that can be sent to arbitration.

It is worth noting that \textit{Circuit City} was decided over a strongly worded dissent joined by three justices. Their strongest argument was that in 1925, when the FAA was passed, Congress’s power under the Commerce Clause was limited, and that Congress never intended to sweep all of these disputes into arbitration.\textsuperscript{351} Given how polarized the Court has been in recent years, and the complete doctrinal reversal detailed above, it is difficult to predict what direction the Supreme Court will take in the future.

2. \textit{Will Sarbanes–Oxley Whistleblowing Claims Be Sent to Arbitration?}

To determine whether Sarbanes–Oxley whistleblowing claims will be sent to arbitration, it is necessary to ask whether the Sarbanes–Oxley Act evinces a Congressional intent to preclude pre-dispute mandatory arbitration contracts. The Supreme Court, in \textit{Gilmer}, established a general presumption in favor of mandatory arbitration.\textsuperscript{352} The decision, however, mentions the text, legislative history, and purposes of an act as ways to rebut that presumption.\textsuperscript{353} Essentially, the \textit{Gilmer} and \textit{Circuit City} line of decisions changed the default rule. Whereas it used to be


\textsuperscript{350} See Stephen J. Ware, \textit{Paying the Price of Process: Judicial Regulation of Consumer Arbitration Agreements}, 2001 J. DISP. RESOL. 89, 89–90 (documenting rise of arbitration clauses in the consumer context and presenting arguments in favor of such clauses).


\textsuperscript{353} \textit{Id.}; see also Shearson/American Express, Inc. v. McMahon, 482 U.S. 220, 227 (1987) (indicating that a party may use text, legislative history, and underlying purposes of a statute to rebut the presumption that an arbitration “agreement” will be enforceable); Rosenberg v. Merrill Lynch, 995 F. Supp. 190, 204 (D. Mass. 1998) (“Instead of overruling \textit{Gilmer} . . . Congress could merely follow [Gilmer’s] . . . suggestion . . . and make clear its intent to preclude enforcement of pre-dispute arbitration . . . .”), aff’d, 170 F.3d 1 (1st Cir. 1999).
assumed that a plaintiff could avail himself or herself of all of the remedies in the statute to the fullest extent, now a presumption in favor of arbitration means that Congress must affirmatively insert language into every act where mandatory arbitration is not to be used.  

The analysis of whether a § 806 claim is subject to arbitration begins with the text of the Sarbanes–Oxley Act, which is silent on the subject of mandatory arbitration. Although one certainly could argue that Congress created its own scheme of enforcement through criminal penalties in § 1107 and civil private rights of action under § 806, neither of which explicitly includes mandatory arbitration, the Supreme Court foreclosed this argument in *Gilmer* and *Circuit City*. Under these decisions, when Congress does not explicitly speak to the question of mandatory arbitration, courts are supposed to apply a presumption in its favor.  

Without express guidance in the text, one must analyze the legislative history. Earlier drafts of the Act included a provision within the whistleblowing section that banned mandatory arbitration. However, those portions of the bill were excised in committee, with no clear rationale emerging for why they were cut. Nevertheless, the committee did eliminate the provision, which does not bode well for a plaintiff arguing for his or her day in court.  

With no express textual support and negative legislative history, the best argument that the Act did not contemplate mandatory arbitration for Sarbanes–Oxley whistleblowing claims must therefore focus on legislative intent. Congress would have been cognizant of the important role of whistleblowers in preventing corporate fraud because Sherron Watkins testified before congressional committees prior to the passage

---

354. See *Gilmer*, 500 U.S. at 26. The burden is on the plaintiff to:

show that Congress intended to preclude a waiver of a judicial forum . . . . If such an intention exists, it will be discoverable in the text[,] . . . its legislative history, or an ‘inherent conflict’ between arbitration and the ADEA’s underlying purposes . . . . Throughout such an inquiry, it should be kept in mind that ‘questions of arbitrability must be addressed with a healthy regard for the federal policy favoring arbitration.

Id. (internal citations omitted).

355. *Id. at 23; Circuit City*, 632 U.S. at 109.


357. An earlier version of the bill contained a provision that stated: “No employee may be compelled to adjudicate his or her rights under this section pursuant to an arbitration agreement.” S. 2010, 107th Cong. § 7 (2002). A bill introduced in the House contained identical language, and would have prevented whistleblowing claims from being sent to arbitration. H.R. 4098, 107th Cong. § 8 (2002).

Employee Whistleblowing After Sarbanes–Oxley

of the Act. Indeed, it was such an important issue to Congress that legislators made retaliation against whistleblowers a crime punishable by ten years imprisonment. Under EEOC v. Waffle House, a government agency can perform investigations and bring an action for violation of a statute, despite the presence of a mandatory arbitration contract or any other private agreement with an individual employee. Therefore, even if courts enforce mandatory arbitration for § 806 claims, the government could still bring a criminal action under § 1107.

In Boss v. Salomon Smith Barney, Inc., the district court examined whether whistleblowing claims under § 806 of Sarbanes–Oxley could be arbitrated. The plaintiff, Kenneth Boss, worked as a research analyst at Smith Barney. Boss alleged that, in violation of the company’s internal rules and procedures, Smith Barney ordered him to give a draft research report to its investment bankers. The investment bankers allegedly urged Boss to change his research findings after receiving the report. Boss alleged that Salomon Smith Barney terminated his employment when he refused to change the report and complained to supervisors. These are serious allegations, because, as noted in Part II.A.1, the Sarbanes–Oxley Act prohibits investment bankers from interfering with the research reports written by analysts within their organizations.

Boss wanted to have his Sarbanes–Oxley whistleblowing claim heard in federal court, but he had signed a mandatory arbitration contract. This contract was one of the routine forms Smith Barney required employees to sign upon joining the firm. Boss had also executed a U-4, the registration form required by the National Association of Securities Dealers (NASD), which requires that employees arbitrate their

359. See Hearings, supra note 26, at 14–67 (testimony of Sherron Watkins).
363. Id. at 684.
364. Id.
365. Id.
366. Id.
367. Id.
370. Id.
disputes.371 Indeed, this was the very form contract at issue in the Gilmer case.372

In a rather conclusory analysis of the issue, the district court granted Smith Barney’s motion to compel arbitration, holding that “[t]here is nothing in the text of the statute or the legislative history of the Sarbanes–Oxley [A]ct evincing intent to preempt arbitration of claims under the act.”373 Further, the court ruled that there was no “‘inherent conflict’ between arbitration and the [statute’s] . . . purposes.”374 Ultimately, the court refused to hear Boss’s whistleblowing claim, and instead sent the case to arbitration with the NASD.375 The court’s disposal of the issue in such a brief fashion flows from the Supreme Court jurisprudence on mandatory arbitration and the failure of Congress to mention arbitration specifically in Sarbanes–Oxley. The court apparently assumed that Congress would be aware of the legal landscape that currently favors arbitration and would also be aware of the fact that, if there was no mention of mandatory arbitration, courts would assume that arbitration was an adequate remedy.376

As I have argued elsewhere, arbitration of employment disputes should be taken with a healthy dose of skepticism, as arbitration in general tends to favor the employer.377 Many arguments, both for (efficiency) and against (fairness), have been raised regarding the use of mandatory arbitration contracts.378 Despite the Supreme Court’s

371. Id. at 685.
374. Id. (quoting Oldroyd v. Elmira Sav. Bank, 134 F.3d 72, 77–78 (2d Cir. 1998)) (alteration in original).
375. Id.
377. See Cherry, supra note 318, at 276–86. One certainty exists: because arbitration results in a reduction of the procedural protections available to employees and does not guarantee a decisionmaker versed in federal employment law, it tends to reduce the settlement value of a plaintiff’s case. Id.
378. For a summary of the various arguments, see id.; Stephen J. Ware, Consumer Arbitration as Exceptional Consumer Law (With a Contractualist Reply to Carrington & Haagen), 29 MCGEORGE
Employee Whistleblowing After Sarbanes–Oxley

imprimatur of approval in *Circuit City*, there is still an argument that arbitration of Sarbanes–Oxley whistleblowing claims should not be foisted on employees at the time of hire, when economic realities dictate that it will be a one-sided negotiation. Indeed, as I have queried elsewhere, if the alternative dispute resolution (ADR) movement values fair process and a mutual meeting of the minds to bargain for a fair solution to underlying goals, then the values of the movement are not consonant with mandatory arbitration. That is because mandatory arbitration is not the result of parties determining that ADR is the best way to resolve a dispute; rather, it is a one-sided provision foisted upon employees through a contract of adhesion.379

An employee can argue the arbitration point in two ways. First, the *Boss* case may have been wrongly decided based on the purposes of the statute and the other protections Sarbanes–Oxley provided to whistleblowers.380 On the other hand, if the *Boss* case was correctly decided because the Supreme Court favors arbitration,381 then Congress made a mistake in not providing whistleblowers with more protection. Either way, a realistic view is that these Sarbanes–Oxley whistleblower cases will be sent to arbitration, as was the case in *Boss*.382 The fact that mandatory arbitration of these claims will be the likely result is a serious weakness in the Act.383

C. Conclusions and Suggested Law Reform

In light of the forgoing discussion, what effect has Sarbanes–Oxley had on corporate governance? What effect will Sarbanes–Oxley have on

---


380. See supra Part II.B.


383. The last apparent weakness in the Act’s treatment of whistleblowers is the allocation of agency oversight of § 806 to OSHA, rather than to the SEC. See *Sarbanes–Oxley Act of 2002*, Pub. L. No. 107-204, § 806, 116 Stat. 745, 802–04 (codified at 18 U.S.C.A. § 1514A (West Supp. 2003)). The SEC is uniformly viewed as the expert agency when it comes to issues of securities violations. In the face of this expertise, giving oversight to OSHA is puzzling. See Deborah Solomon, *For Financial Whistle-Blowers, New Shield Is an Imperfect One*, WALL ST. J., Oct. 4, 2004, at A1 (pointing out OSHA’s lack of expertise in accounting matters as well as its limited authority to subpoena witnesses). However, because OSHA does have experience with workplace disputes and other whistleblowing statutes, the question of OSHA’s effectiveness is an open one.
the Sherron Watkinses and Cynthia Cooper of the world? Unfortunately, the answer to both of these questions seems to be “not as much as it should.” The promise of Sarbanes–Oxley is its seeming federalization of state retaliatory discharge law regarding accounting fraud, but its weakness is in the remedies provided.

Employers must take action to deal with employee complaints of the type that both Cooper and Watkins made, and yet the precise mode of dealing with such complaints is not specified in the Act. Information can seemingly be sent into a void, and yet an employer will still comply with the letter of the law. Further, whistleblowing complaints under Sarbanes–Oxley can be sent to arbitration, a forum that provides inadequate remedies for employees. Some retaliatory actions are now considered criminal—even though the underlying employment action could be sent to mandatory arbitration. In short, the Sarbanes–Oxley Act provides a strange combination of tort, criminal, employment, and alternative dispute resolution issues in its reform of federal securities laws.

If Sarbanes–Oxley had been in force at the time of the scandals, what would have been the fate of the two whistleblowers, Watkins and Cooper? Those at the top of the Enron hierarchy could still have asked outside counsel about the company’s ability to fire Sherron Watkins, but after Sarbanes–Oxley, the answer would be that Enron could not fire Watkins for reporting her concerns.384 That answer would be definitive and would no longer depend on the vagaries of the Texas public policy exception. If Watkins had been fired or another adverse employment action taken against her for making such a report, Enron could have faced a lawsuit for retaliatory discharge under § 806.385 The suit would likely have resulted in Enron providing Watkins with “make whole” relief.386 Further, after her testimony, any participants in her firing could have faced criminal charges under § 1107.387

In that respect, Sarbanes–Oxley is an advance for conscientious employees, and a further movement away from the at-will employment rule. However, § 806 is an area-specific whistleblowing statute; it applies only to fraud.388 There are still many other areas where

385. Id.
386. Id.
387. Id. § 1107 (codified at 18 U.S.C.A. § 1513 (West Supp. 2003)).
388. Id. § 806 (codified at 18 U.S.C.A. § 1514A (West Supp. 2003)).
employees may feel compelled to report violations of a state or federal rule or regulation, and such reports would be in the public interest, yet employees could be dismissed from their positions, harassed, or otherwise retaliated against. The survey in Appendix A indicates that in many jurisdictions without broad whistleblower protection statutes, or where judicial decisions have given such statutes narrow application, it is relatively easy for an employee’s whistleblowing claim to fall through the cracks.

Law reform, then, should include uniform federal protection for whistleblowers who report any violation of federal law or regulation to law enforcement. This is with the caveat that there should be an “escape clause” for such things as reporting that someone in the company was involved in an essentially technical violation. This protection could cover workplaces over a certain size—perhaps fifteen, as is the case with Title VII—and would not depend on either the vagaries of the type of wrongdoing that is being reported or the jurisdiction in which the whistleblower happens to live. Such protection would have a positive impact not only on the individual whistleblowers who receive direct protection under the law, but also on the enforcement of federal statutes. A uniform statement of state law, standardizing the types of dismissal that are against public policy, would also be a positive development.

In conclusion, Sarbanes–Oxley will have an impact on employment law. The Act federalizes (and, indeed, potentially criminalizes) an entire swath of activity that (depending upon state statutes and common law) might not have even been actionable as a civil tort under the previous legal regime. On the other hand, the Act does not provide adequate remedies, which provokes the question whether the Act will result in the right mix of incentives and disincentives for the optimal amount of whistleblowing.

Part of the whistleblower’s decision depends on the incentives or disincentives the law provides. Worker behavior tends to be influenced by the applicable law. However, whistleblowing behavior is partially a matter of individual conscience. Would the passage of Sarbanes–Oxley have influenced Watkins’s decision to write her letter? She decided to make her report, knowing that she faced firing, demotion, or other forms
of retaliation—knowing that there would likely be no recourse or remedy in law against Enron. In a way, it seems that Watkins, like many whistleblowers, felt an obligation to make a report regardless of any protection that the legal system would afford her. At the same time, those who are on the margins, those employees who are trying to decide if reporting illegal activity is worth losing their jobs, may be influenced if they know that retaliatory discharges are against the law.

Ultimately, the law should support those who are willing to make personal sacrifices for the public good and increase support for those who blow the whistle. The law should protect whistleblowers so that the high social price they have to pay is eased, not compounded. Cynthia Cooper recently gave some advice to a graduating class from her alma mater, Mississippi State University. Her speech focused on an employee’s sense of honor, and her comment to students was that they should “[do] what you know is right even if there may be a price to be paid.” Employment law doctrine should provide full support for such ethics and integrity.

392. Interview with Sherron Watkins in Birmingham, Ala. (Feb. 18, 2004).
393. Scott Waller, ‘Do what . . . is right’: Whistleblower Tells Students to Have Personal Integrity, CLARION-LEDGER (Jackson, Miss.), Nov. 18, 2003, at C1.
APPENDIX A

This Appendix summarizes, by state, the applicable causes of action for whistleblowing, listing and briefly describing every state whistleblowing statute that is potentially relevant to disclosure of information by employees. In addition, this Appendix briefly lists relevant caselaw from each jurisdiction and any pertinent secondary material particular to that state. This Appendix owes a debt to Callahan and Dworkin’s survey of fifty states, supra note 80, at 107–08.

ALABAMA: Alabama does not recognize a public policy exception for private employees. See, e.g., Salter v. Alfa Ins. Co., 561 So. 2d 1050, 1053 (Ala. 1990); Hinrichs v. Tranquilaire Hosp., 352 So. 2d 1130, 1131 (Ala. 1977). However, state employees are protected if the employee reports “under oath or in the form of an affidavit, a violation of a law, a regulation, or a rule . . . to a public body” or reports a code of ethics violation. ALA. CODE § 36-26A-3 (2001); id. § 36-25-24 (providing protection to state employees who report code of ethics violation); see State Employees Protection Act, id. §§ 36-26A-1 to -7.


ALASKA: Alaska recognizes a public policy exception for private and public employees, as well as an implied covenant of good faith and fair dealing. See Reed v. Municipality of Anchorage, 782 P.2d 1155, 1158 (Alaska 1989); Knight v. Am. Guard & Alert, Inc., 714 P.2d 788, 792 (Alaska 1986). The Alaska Whistleblower Act protects a public employee who “reports to a public body or is about to report to a public body a matter of public concern” or “participates in a court action, an investigation, a hearing, or an inquiry held by a public body on a matter of public concern.” Alaska Whistleblower Act, ALASKA STAT. §§ 39.90.100–.150 (Michie 2002 & Supp. 2003); see also ALASKA STAT. § 23.40.110 (Michie 2002) (unfair labor practices—public
Area-specific whistleblower statutes: **Alaska**:

- **Alaska Stat. § 18.60.089** (occupational safety and health); *id.* § 18.80.220 (unlawful employment practice—discrimination); **Alaska Stat. § 23.10.135** (Wage and Hour Act); **Alaska Stat. § 24.60.035** (reports to legislative ethics committee or other government entity); **Alaska Stat. § 39.28.070** (affirmative action or equal employment opportunity); **Alaska Stat. § 42.40.760** (unfair labor practices—Alaska Railroad Corporation); **Alaska Stat. § 47.24.120** (vulnerable adults); *id.* § 47.35.105 (social services for children, maternity homes); *id.* § 47.62.040 (long-term care).


Area-specific whistleblower statutes: **Arizona**:

- **ARIZ. REV. STAT. § 3-376** (pesticide control); *id.* § 3-3120 (occupational safety and health—agricultural employment); **ARIZ. REV. STAT. § 23-329** (minimum wages for minors); *id.* § 23-425 (occupational safety and health); *id.* § 23-1385 (unfair labor practices—agricultural employees); **ARIZ. REV. STAT. § 36-2282** (health care facility employees—infant care); **ARIZ. REV. STAT. § 41-1378** (ombudsman-citizens aide); *id.* § 41-1464 (discrimination in employment); *id.* § 41-1492.10 (public accommodation and services); **ARIZ. REV. STAT. §§ 49-207, 49-
Employee Whistleblowing After Sarbanes–Oxley


Area-specific whistleblower statutes: ARK. CODE ANN. § 5-55-113 (Michie 1987) (Medicaid fraud—provision for reward); ARK. CODE ANN. § 9-28-405(f)(1) (Michie 1987) (retaliation grounds for revocation/denial of child welfare agency license); ARK. CODE ANN. § 11-3-204 (Michie 1987) (references to prospective employers—retaliation causes loss of immunity for employers); id. § 11-4-206 (minimum wage law); id. § 11-4-608 (wage discrimination); id. § 11-10-106 (employment security); ARK. CODE ANN. § 16-123-108 (Michie 1987 & Supp. 2003) (Arkansas Civil Rights Act of 1993); id. § 16-123-208 (fair housing); ARK. CODE ANN. § 20-10-1007 (Michie 1987) (long-term care facilities and services).

CALIFORNIA: California recognizes a public policy exception for at-will employees, often referred to as a Tameny claim. See Green v. Ralee Eng’g Co., 960 P.2d 1046, 1051 (Cal. 1998); Tameny v. Atl. Richfield Co., 610 P.2d 1330, 1335–37 (Cal. 1980). Private and public employees are protected if the employee reports violations of state or federal statutes and regulations to “a government or law enforcement agency.” CAL. LAB. CODE § 1102.5 (West 2003). State employees are also protected if reporting “an improper governmental activity or . . . any condition that may significantly threaten the health or safety of employees or the public . . . .” CAL. GOV’T CODE § 8547.2 (West 1992 & Supp. 2004); see California Whistleblower Protection Act, id. §§ 8547–8547.12; see also CAL. GOV’T CODE § 19702 (West 1995 & Supp. 2004) (discrimination in public employment); CAL. GOV’T CODE
§ 53298 (West 1997) (local government employees).

**Area-specific whistleblower statutes:** CAL. BUS. & PROF. CODE § 2056 (West 2003) (physicians advocating for medically appropriate health care); CAL. BUS. & PROF. CODE § 7583.46 (West Supp. 2004) (private patrol operator); CAL. BUS. & PROF. CODE § 9998.6 (West 1995) (foreign workers); CAL. EDUC. CODE § 44040 (West 1993) (school district employees—appearance before certain boards or committees); CAL. EDUC. CODE § 44114 (West 1993 & Supp. 2004) (school employees—improper governmental activities); CAL. EDUC. CODE § 87039 (West 2002) (community college district employee—appearance before certain boards or committees); id. § 87164 (community college employees—improper governmental activities); CAL. FIN. CODE § 6530 (West 1999) (savings associations); CAL. GOV’T CODE §§ 9149.20–9149.23 (West Supp. 2004) (disclosure of improper governmental activities to legislative committee); CAL. GOV’T CODE § 12653 (West 1992) (false claims actions); id. § 12940 (unlawful employment practices—discrimination); id. § 12944 (retaliation by licensing board); id. § 12945.2 (family care leave); id. § 19251.5 (communication with Legislature); CAL. HEALTH & SAFETY CODE § 444.22 (West 1990 & Supp. 2004) (communication with consumer health hotlines); CAL. HEALTH & SAFETY CODE § 1278.5 (West 2000) (care, services, or conditions of health care facilities); id. § 1317.4 (report of violations—health facilities); CAL. HEALTH & SAFETY CODE § 1432 (West 2000 & Supp. 2004) (long-term health care facilities); CAL. HEALTH & SAFETY CODE § 1539 (West 2000) (inspection of community care facilities); id. § 1568.07 (inspection of residential care facilities for persons with chronic life-threatening illnesses); id. § 1569.37 (inspection of residential care facilities for the elderly); id. § 1581 (inspection of adult day health care centers); id. § 1596.881 (child day care facilities); CAL. HEALTH & SAFETY CODE § 17031.7 (West 1984) (employee community housing); CAL. INS. CODE § 1871.7 (West 1993 & Supp. 2004) (fraudulent insurance claims action); CAL. LAB. CODE § 98.6 (West 2003) (exercise of employee rights); id. § 132a (testimony in co-employee’s workers’ compensation case); CAL. LAB. CODE § 752 (West 2003 & Supp. 2004) (complaint regarding election—smelters and underground workings); CAL. LAB. CODE § 1153 (West 2003) (unfair labor practices—agricultural employers); id. § 6310 (occupational safety and health); id. § 6399.7 (hazardous substances information and training); CAL. MIL. & VET. CODE § 73.7 (West Supp. 2004) (veterans programs); CAL. PENAL CODE § 6129 (West 2000 &
Employee Whistleblowing After Sarbanes–Oxley


COLORADO: Colorado recognizes a public policy exception to the at-will employment rule. See Flores v. Am. Pharm. Servs., Inc., 994 P.2d 455, 458–59 (Colo. 1999). State employees are protected when disclosing information “regarding any action, policy, regulation, practice, or procedure, including, but not limited to, the waste of public funds, abuse of authority, or mismanagement of any state agency.” COLO. REV. STAT. ANN. § 24-50.5-102 (West 2003); see id. §§ 24-50.5-101 to 24-50.5-107 (state employee protection).

Area-specific whistleblower statutes: COLO. REV. STAT. ANN. § 8-2.5-101 (West 2003) (communication with general assembly or court of law); id. § 8-3-108 (unfair labor practices—Labor Peace Act); id. § 8-4-120 (wages); id. § 8-6-115 (minimum wage); COLO. REV. STAT. ANN. § 24-34-402 (West 2003) (discriminatory or unfair employment practices); id. § 24-114-102 (private enterprise employee protection); COLO. REV. STAT. ANN. § 26-3.1-102 (West 2003) (abuse of at-risk adult); id. § 26-3.1-204 (financial exploitation of at-risk adult); id. § 26-11.5-109 (long-term care); COLO. REV. STAT. ANN. § 28-3.1-604 (West 2003) (state military forces—complaints of wrongs).


CONNECTICUT: Connecticut recognizes a public policy exception to the at-will employment rule. See Sheets v. Teddy’s Frosted Foods, Inc., 427 A.2d 385, 388–89 (Conn. 1980). Public and private employees are protected if the employee reports violations of laws or regulations to a public body. CONN. GEN. STAT. § 31-51m (2003). State employees (and employees of large state contractors) are also protected if the employee provides information to the Auditor of Public Accounts or the state Attorney General. CONN. GEN. STAT. § 4-61dd (2003); see also CONN. GEN. STAT. § 5-272 (2003) (collective bargaining—state employees); CONN. GEN. STAT. § 7-470 (2003) (collective bargaining—municipal employees).
Area-specific whistleblower statutes: CONN. GEN. STAT. § 4-37j (2003) (whistleblower protection for foundation employees); CONN. GEN. STAT. § 10-153e (2003) (certified professional employees—education); CONN. GEN. STAT. § 16-8a (2003) (public service companies); CONN. GEN. STAT. § 17a-101e (2003) (child abuse or neglect); id. § 17b-410 (long-term care); CONN. GEN. STAT. § 19a-532 (2003) (nursing home facilities); CONN. GEN. STAT. § 31-40d (2003) (employers who use or produce carcinogens); id. § 31-40t (hazardous conditions); id. § 31-51pp (family and medical leave); id. § 31-69 (testimony before wage board); id. § 31-69b (wages; state contracts); id. § 31-76 (wage discrimination); id. § 31-105 (unfair labor practices—Labor Relations Act); id. § 31-226a (unemployment compensation); id. § 31-379 (occupational safety and health); CONN. GEN. STAT. § 46a-13e (Victim Advocate); id. § 46a-13n (Child Advocate); id. § 46a-60 (discriminatory employment practice); CONN. GEN. STAT. § 54-85b (2003) (testimony in any criminal proceeding); id. § 54-203 (victim appearing as witness in any criminal proceeding).


Employee Whistleblowing After Sarbanes–Oxley


FLORIDA: Florida courts have refused to recognize a public policy exception to the at-will employment rule. The courts view the creation of such protection as the proper province of the legislature. Hartley v. Ocean Reef Club, Inc., 476 So. 2d 1327, 1329–30 (Fla. Dist. Ct. App. 1985). Legislative response includes a Florida whistleblower statute, which provides protection for all employees, public and private, who report violations of law. See FLA. STAT. ANN. § 112.3187 (West 2002 & Supp. 2004) (violations by public agency); FLA. STAT. ANN. §§ 448.101–105 (West 2002) (protection for employee reporting violation of law, rule, or regulation by private employer).

Area-specific whistleblower statutes: FLA. STAT. ANN. § 39.203 (West 2003) (child abuse—facilities serving children); FLA. STAT. ANN. § 68.088 (West 1997) (False Claims Act); FLA. STAT. ANN. § 92.57 (West 1999) (testimony in judicial proceeding); FLA. STAT. ANN.


**GEORGIA:** Georgia does not recognize a public policy exception for private employees. See Eckhardt v. Yerkes Reg’l Primate Ctr., 561 S.E.2d 164, 165–66 (Ga. Ct. App. 2002) (holding that there is no public policy exception in Georgia and refusing to allow claim where plaintiffs were fired for reporting that monkeys were infected with deadly disease potentially communicable to humans). However, state employees are protected if they report fraud, waste, or abuse “unless the complaint was made . . . with the knowledge that it was false or with willful disregard for its truth or falsity.” GA. CODE ANN. § 45-1-4 (2002).


**HAWAII:** Hawaii recognizes a public policy exception to the at-will employment rule. Parnar v. Americana Hotels, Inc., 652 P.2d 625, 631 (Haw. 1982). Public and private employees are protected if the employee reports to the employer or a public body violations of law or of contracts “executed by the State, a political subdivision of the State, or the United States . . . .” HAW. REV. STAT. § 378-62 (2003); see Whistleblowers’
Employee Whistleblowing After Sarbanes–Oxley

Protection Act, id. §§ 378-61 to 378-69; see also HAW. REV. STAT. § 89-13 (2003) (prohibited practices—collective bargaining in public employment).


**IDAHO:** Idaho recognizes a public policy exception to the at-will employment rule. See Thomas v. Med. Ctr. Physicians, P.A., 61 P.3d 557, 565 (Idaho 2002). Public employees are protected if the employee reports, in good faith, a violation of law or “any waste of public funds, property or manpower . . . .” IDAHO CODE § 6-2104 (Michie 2004); see Idaho Protection of Public Employees Act, id. §§ 6-2101 to 6-2109.

**Area-specific whistleblower statutes:** IDAHO CODE § 39-4420 (Michie 2002) (hazardous waste management); id. § 39-6214 (PCB waste disposal); IDAHO CODE § 44-1509 (Michie 2003) (minimum wage); id. § 44-1615 (farm labor contractors); id. § 44-1702 (wage discrimination based upon sex); id. § 44-1904 (sanitation facilities for farm workers); IDAHO CODE § 45-613 (Michie 2003) (claims for wages); IDAHO CODE § 67-5009 (Michie 2001) (long-term care); id. § 67-5911 (human rights).


**ILLINOIS:** Illinois recognizes a public policy exception to the at-will employment rule. See Palmateer v. Int’l Harvester Co., 421 N.E.2d 876, 879–80 (Ill. 1981). Private employees are protected if the employee reports to a “government or law enforcement agency . . . a violation of a State or federal law, rule, or regulation.” 740 ILL. COMP. STAT. ANN.
174/15 (West Supp. 2004); see Whistleblower Act, id. 174/1–35 (West Supp. 2004). Private citizens may bring a qui tam action in the event of fraud against the government. 740 ILL. COMP. STAT. ANN. 175/1–8 (Whistleblower Reward and Protection Act) (West 1993 & Supp. 2004); see also 5 ILL. COMP. STAT. ANN. 315/10 (unfair labor practices—Illinois Public Labor Relations Act); 20 ILL. COMP. STAT. ANN. 415/19c.1 (West 2001) (Executive branch—disclosure of prohibited activity).

**Area-specific whistleblower statutes:** 20 ILL. COMP. STAT. ANN. 105/4.04 (West 2001) (long-term care facilities); 105 ILL. COMP. STAT. ANN. 5/34-2.4c (West 1998) (board of education employees); 115 ILL. COMP. STAT. ANN. 5/14 (West 1998) (unfair labor practices—educational employees); 210 ILL. COMP. STAT. ANN. 35/10 (West 2000) (Community Living Facilities); id. 45/3-318 (Nursing Home Care Act); id. 45/3-608 (nursing homes); 210 ILL. COMP. STAT. ANN. 85/10.8 (West Supp. 2004) (employment of physicians); id. 86/35 (hospital employers); 210 ILL. COMP. STAT. ANN. 105/13 (West 2000) (agricultural workers); id. 110/17 (migrant workers); 225 ILL. COMP. STAT. ANN. 10/7.2 (West 1998) (employees of licensees subject to Child Care Act of 1969); id. 705/13.13 (Coal Mining Act); 320 ILL. COMP. STAT. ANN. 20/4.1 (West 2001) (elder abuse and neglect); 325 ILL. COMP. STAT. ANN. 5/9.1 (West 2001) (child abuse and neglect); 415 ILL. COMP. STAT. ANN. 5/52 (West 1997) (Environmental Protection Act); 740 ILL. COMP. STAT. ANN. 92/40 (West 2002) (Insurance Claims Fraud Prevention Act); 740 ILL. COMP. STAT. ANN. 175/1–8 (West 2002 & Supp. 2004) (Whistleblower Reward and Protection Act); 775 ILL. COMP. STAT. ANN. 5/6-101 (West 2001) (civil rights violations); 820 ILL. COMP. STAT. ANN. 40/12 (West 1999) (Personnel Record Review Act); id. 55/15 (Privacy in the Workplace Act); 820 ILL. COMP. STAT. ANN. id. 105/11 (minimum wage); 820 ILL. COMP. STAT. ANN. 112/10, /35 (West Supp. 2004); 820 ILL. COMP. STAT. ANN. 115/14 (West 1999) (wage payment and collection); id. 125/15 (testimony before wage board); id. 130/11b (Prevailing Wage Act); 820 ILL. COMP. STAT. ANN. 180/20 (West Supp. 2004) (leave due to domestic or sexual violence); 820 ILL. COMP. STAT. ANN. 220/2 (West 1993) (Safety Inspection and Education Act); id. 255/14 (Toxic Substances Disclosure to Employees Act).

Employee Whistleblowing After Sarbanes-Oxley


**Area-specific whistleblower statutes:** IND. CODE ANN. § 4-2-6-13 (Michie 2002) (state ethics commission); id. § 4-13-1.2-11 (department of correction ombudsman); IND. CODE ANN. § 5-11-1-9.5 (Michie 2001) (sworn statements to state examiner); IND. CODE ANN. § 12-10-3-11 (Michie 2001) (report of endangered adult); id. § 12-10-13-20 (long-term care); id. § 12-11-13-16 (state wide waiver ombudsman); IND. CODE ANN. § 16-28-9-3 (Michie 1998) (health care facilities); IND. CODE ANN. § 16-41-11-8 (Michie 1993) (training in health precautions for communicable diseases); IND. CODE ANN. § 20-7.5-1-7 (Michie 1997) (unfair practices—school employees); IND. CODE ANN. § 20-12-1-8 (Michie 2000) (state educational institution employees); IND. CODE ANN. § 22-5-3-3 (Michie 1997) (private employer under public contract); id. § 22-8-1.1-38.1 (Occupational Health & Safety); id. § 22-9-1-6 (information provided to Civil Rights Commission).

**IOWA:** Iowa recognizes a public policy exception to the at-will employment rule. See Fitzgerald v. Salsbury Chem., Inc., 613 N.W.2d 275, 281–85 (Iowa 2000). State employees are protected if an employee reports “a violation of law or rule, mismanagement, a gross abuse of funds, an abuse of authority, or a substantial and specific danger to public health or safety.” IOWA CODE ANN. § 70A.28 (West 1999 & Supp. 2004); see also IOWA CODE ANN. § 19A.19 (West 2001) (state personnel); IOWA CODE ANN. § 20.10 (West 2001) (Public Employment Relations); IOWA CODE ANN. § 70A.29 (West 1999) (employees of political subdivisions).

**Area-specific whistleblower statutes:** IOWA CODE ANN. § 68B.32A (West 1999) (information provided to Ethics & Campaign Disclosure Board); IOWA CODE ANN. § 88.9 (West 1996) (occupational safety and health); IOWA CODE ANN. § 89B.9 (West 1996) (hazardous chemical right to know); IOWA CODE ANN. § 91A.10 (West 1996) (wage payment collection); IOWA CODE ANN. § 135C.46 (West 1997) (health care facilities); IOWA CODE ANN. § 207.28 (West 2000) (coal mining); IOWA
KANSAS: Kansas recognizes a public policy exception to the at-will employment rule. See Palmer v. Brown, 752 P.2d 685, 689–90 (Kan. 1988). State agency employees are protected if the employee reports a violation of a state or federal laws, rules, or regulations. KAN. STAT. ANN. § 75-2973 (1997 & Supp. 2003); see also KAN. STAT. ANN. § 75-4333 (1997) (public employer and employee relations).

Area-specific whistleblower statutes: KAN. STAT. ANN. § 38-1525 (2000) (child abuse); KAN. STAT. ANN. § 39-1403 (2000) (dependent adult abuse); id. § 39-1432 (dependent adult abuse); KAN. STAT. ANN. § 44-615 (2000) (information provided to secretary of human resources); id. § 44-636 (unsafe or hazardous working conditions); id. § 44-808 (employer and employee relations); id. § 44-828 (agricultural labor relations); id. § 44-1009 (discrimination); id. § 44-1113 (age discrimination); id. § 44-1210 (wages and hours); KAN. STAT. ANN. § 47-852 (2000) (impaired veterinarians); KAN. STAT. ANN. § 65-4928 (2002) (health care—incidents and impaired health care providers); KAN. STAT. ANN. § 72-5430 (2002) (retaliation by board of education against professional employee); id. § 72-89b04 (school employees reporting criminal acts); KAN. STAT. ANN. § 75-7313 (Supp. 2003) (long-term care).


KENTUCKY: Kentucky recognizes a public policy exception to the at-will employment rule. See Grzyb v. Evans, 700 S.W.2d 399, 401–02 (Ky. 1985). Public employees are protected if they report a violation of law or “mismanagement, waste, fraud, abuse of authority, or a substantial and specific danger to public health or safety.” KY. REV. STAT. ANN. § 61.102 (Banks–Baldwin 2001).

Area-specific whistleblower statutes: KY. REV. STAT. ANN. § 6.915 (Banks–Baldwin 1997) (information provided to Legislative Program Review and Investigations Committee); KY. REV. STAT. ANN. § 205.8465 (Banks–Baldwin 2001) (public and medical assistance fraud and abuse); KY. REV. STAT. ANN. § 207.170 (Banks–Baldwin 2001)
Employee Whistleblowing After Sarbanes–Oxley


Case law: Palmer v. Int’l Ass’n of Machinists, 882 S.W.2d 117, 120–21 (Ky. 1994).


MAINE: Maine does not recognize a public policy exception for private employees. See Bard v. Bath Iron Works, 590 A.2d 152, 156 (Me. 1991) ("[T]his court has yet to recognize a common law action for wrongful discharge."); Larrabee v. Penobscot Frozen Foods, Inc., 486 A.2d 97, 100 (Me. 1984); Lyons v. La. Pac. Corp., No. 02-29-B-K, 2002 U.S. Dist. LEXIS 6042, at *9 (D. Me. Apr. 5, 2002). Employees, both public and private, are protected if the employee reports a violation of a law or rule or a “condition or practice that would put at risk the health or safety of that employee or any other individual.” ME. REV. STAT. ANN. tit. 26, § 833 (West 1988 & Supp. 2003); see Whistleblowers’ Protection Act, id. §§ 831–840; see also ME. REV. STAT. ANN. tit. 26, § 964 (West 1988) (Municipal Public Employees Labor Relations Law); ME. REV. STAT. ANN. tit. 26, § 979-C (West 1988 & Supp. 2003) (State Employees Labor Relations Act); id. § 1027 (University of Maine System Labor Relations Act); ME. REV. STAT. ANN. tit. 26, § 1284 (West 1988) (Judicial Employees Labor Relations Act).


MARYLAND: Maryland recognizes a public policy exception to the at-will employment rule. See Wholey v. Sears Roebuck, 803 A.2d 482, 494 (Md. 2002). Executive branch employees are protected if the employee reports a violation of law, “an abuse of authority, gross mismanagement, or gross waste of money [or] . . . a substantial and specific danger to public health or safety.” MD. CODE ANN., STATE
Employee Whistleblowing After Sarbanes–Oxley

PERS. & PENS. § 5-305 (1999 & Supp. 2003); see id. §§ 5-301 to -313 (Executive Branch employees); see also MD. CODE ANN., STATE PERS. & PENS. § 2-305 (1997) (state employment).


MASSACHUSETTS: Massachusetts recognizes a public policy exception to the at-will employment rule. See Shea v. Emmanuel Col., 682 N.E.2d 1348, 1349–50 (Mass. 1997). Public employees are protected if the employee reports a violation of law or “a risk to public health, safety or the environment.” MASS. ANN. LAWS ch. 149, § 185 (Law. Co-op. 1999); see also id. ch. 150E, § 10 (public employees— labor relations).


**MICHIGAN:** Although Michigan recognizes a public policy exception to the at-will employment rule, it applies only to a limited number of cases. See Dudewicz v. Norris-Schmid, Inc., 503 N.W.2d 645, 649–50 (Mich. 1993). Michigan courts have recognized a public policy cause of action where the statute that forms the basis for the public policy at issue does not provide a remedy for retaliatory discharge. Id.; see also Trombetta v. Detroit, Toledo & Ironton R.R. Co., 265 N.W.2d 385, 388 (Mich. Ct. App. 1978). Thus, a plaintiff reporting a violation of law cannot maintain a cause of action because the plaintiff has a remedy under the Whistleblowers’ Protection Act (WPA). See Dudewicz, 503 N.W.2d at 650. However, in a case decided prior to the enactment of the WPA, the Michigan State Court of Appeals held that the plaintiff had stated a cause of action for retaliatory discharge based on his refusal to falsify reports, which would constitute a violation of law. Trombetta, 265 N.W.2d at 388. The Michigan whistleblower statute applies to both private and public employers, as the statute defines “employer” broadly. See Whistleblowers’ Protection Act, MICH. COMP. LAWS ANN. §§ 15.361–369 (West 1994).

**Area-specific statutes:** MICH. COMP. LAWS ANN. § 18.1486 (West 1994 & Supp. 2003) (internal auditors); MICH. COMP. LAWS ANN. § 37.1602 (West 2001) (Persons With Disabilities Civil Rights Act); id. § 37.2701 (civil rights); MICH. COMP. LAWS ANN. § 259.114 (West Supp. 2004) (audit committee and internal auditor—Airport Authorities); MICH. COMP. LAWS ANN. §§ 324.6316, 324.6532 (West 1999) (vehicle emissions testing); MICH. COMP. LAWS ANN. § 330.1755 (West 1999) (mental health services); MICH. COMP. LAWS ANN. § 333.16244 (West 2001) (health professional organization employees); id. § 333.20176a (medical malpractice); MICH. COMP. LAWS ANN. § 333.20180 (West 2001 & Supp. 2003) (employees/contractors of health facilities and agencies); MICH. COMP. LAWS ANN. § 333.21771
Employee Whistleblowing After Sarbanes–Oxley

(West 2001) (nursing homes); MICH. COMP. LAWS ANN. § 400.586j (West 1997) (long-term care facilities); MICH. COMP. LAWS ANN. § 408.395 (West 1999) (minimum wage); id. § 408.483 (payment of wages and fringe benefits); id. §§ 408.1029, 408.1065 (occupational safety and health); MICH. COMP. LAWS ANN. § 423.16 (West 2001) (labor disputes—mediation); id. § 423.210 (labor disputes—strikes by public employees); MICH. COMP. LAWS ANN. § 750.145p (West Supp. 2003) (vulnerable adults).


MISSISSIPPI: The Supreme Court of Mississippi adopted a public policy exception to the at-will employment rule in 1993. See McArn v. Allied Bruce-Terminix Co., 626 So. 2d 603, 607 (Miss. 1993).


MISSOURI: The Missouri Court of Appeals has allowed a public policy exception to the at-will employment rule. See Luethans v. Wash. Univ., 894 S.W.2d 169, 171 n.2 (Mo. 1995); Boyle v. Vista Eyewear, Inc., 700 S.W.2d 859, 870–77 (Mo. Ct. App. 1985). State employees are protected if they report a violation of law or “[m]ismanagement, a gross waste of funds or abuse of authority, or a substantial and specific danger to public health or safety . . . .” MO. ANN. STAT. § 105.055 (West 2003).

Area-specific whistleblower statutes: MO. ANN. STAT. § 105.467 (West 2003) (conflicts of interest—lobbying); MO. ANN. STAT. § 188.105 (West 2003) (complaint/testimony/assistance—regulation of abortions); MO. ANN. STAT. § 197.285 (West 2003) (hospital and ambulatory surgical center employees); MO. ANN. STAT. § 198.070 (West 2003) (abuse—nursing homes); id. § 198.090 (misappropriation of funds—nursing homes); MO. ANN. STAT. § 213.070 (West 2003) (human rights); MO. ANN. STAT. § 217.410 (West 2003) (inmate abuse); MO. ANN. STAT. § 290.525 (West 2003) (minimum wage); MO. ANN. STAT. § 630.167 (West 2003) (mental health facilities); id. § 630.635 (testimony/information—mental health placement); MO. ANN. STAT. § 633.135 (West 2003) (testimony/information—mental health placement or discharge); MO. ANN. STAT. § 660.300 (West 2003) (abuse and neglect—in-home services); id. § 660.305 (misappropriation of property—in-home services); id. § 660.608 (long-term care).

Case law: Bachtel v. Miller County Nursing Home Dist., 110 S.W.3d 799 (Mo. 2003).


MONTANA: Montana is the only state to have abolished the at-will
Employee Whistleblowing After Sarbanes–Oxley

rule. It has a wrongful-discharge statute that requires an employer to show “good cause” for termination. Wrongful Discharge From Employment Act, MONT. CODE ANN. §§ 39-2-901 to -915 (2003).

**Area-specific whistleblowing statutes:** MONT. CODE ANN. § 2-18-1001 (2003) (Transportation Department personnel filing grievances); id. § 2-18-1011 (filing grievances—hearing on complaint); MONT. CODE ANN. § 39-31-401 (2003) (unfair labor practices—public employees); MONT. CODE ANN. § 49-2-301 (2003) (prohibited discriminatory practices); id. § 49-3-209(4) (governmental code of fair practices); MONT. CODE ANN. § 50-78-204 (2003) (hazardous chemical exposure—employee rights). However, these are seemingly irrelevant because of the good cause requirement.


**NEBRASKA:** Nebraska recognizes a public policy exception to the at-will employment rule. See Schriner v. Meginnis Ford Co., 421 N.W.2d 755, 757–59 (Neb. 1988). The State Government Effectiveness Act protects public employees who report a violation of law, “gross mismanagement or gross waste of funds, or . . . a substantial and specific danger to public health or safety.” NEB. REV. STAT. § 81-2703 (1999); see id. § 81-2705; see also id. § 81-1386 (collective bargaining—state employees).


**NEVADA:** Nevada recognizes a public policy exception to the at-will employment rule. See Allum v. Valley Bank, 970 P.2d 1062, 1063–64 (Nev. 1998). State and local government employees are protected if they disclose improper governmental action, which includes violation of state law or regulation, “abuse of authority[,] . . . substantial and specific danger to the public health or safety[,] or . . . gross waste of public


Employee Whistleblowing After Sarbanes–Oxley

**Case law:** Appeal of Fred Fuller Oil Co., 744 A.2d 1141 (N.H. 2000).


**NEW MEXICO:** New Mexico recognizes a public policy exception to the at-will employment rule. *See* Silva v. Am. Fed’n of State, County & Mun. Employees, 37 P.3d 81, 83 (N.M. 2001).

**Area-specific whistleblower statutes:** N.M. STAT. ANN. § 10-7E-19(E) (Michie 2003) (Public Employee Bargaining Act); N.M. STAT. ANN. § 22-10A-33(B) (Michie 2003) (school employees—violence and vandalism); N.M. STAT. ANN. § 28-1-7 (Michie 2000 &
NEW YORK: New York does not recognize a public policy exception to the at-will employment rule. Horn v. N.Y. Times, 790 N.E.2d 753, 759 (N.Y. 2003). A private employee is protected if the employee “discloses, or threatens to disclose to a supervisor or to a public body an activity, policy or practice of the employer that is in violation of law, rule or regulation which violation creates and presents a substantial and specific danger to the public health or safety . . . .” N.Y. LAB. LAW § 740 (McKinney 2002). To maintain an action under section 740, reasonable belief of a violation of law is insufficient; proof of an actual violation is required. Bordell v. Gen. Elec. Co., 667 N.E.2d 922, 923 (N.Y. 1996). Public employees are protected if they report a “violation of a law, rule or regulation which violation creates and presents a substantial and specific danger to the public health or safety; or . . . which the employee reasonably believes to be true and reasonably believes constitutes an improper governmental action.” N.Y. CIV. SERV. LAW § 75-b(2)(a) (McKinney 1999); see also N.Y. GEN. MUN. LAW § 683 (McKinney 1999) (public employees—grievances).

Area-specific whistle blower statutes: N.Y. EDUC. LAW § 3028-c (McKinney 2001) (school employees—acts of violence and weapons possession); N.Y. EXEC. LAW § 296 (McKinney 2001 & Supp. 2004) (unlawful discriminatory practices—human rights); N.Y. EXEC. LAW § 544-a(15)(c) (McKinney 1996 & Supp. 2004) (long-term care facilities); N.Y. LAB. LAW § 27-a (McKinney 2002) (safety and health standards); id. § 215(1) (labor law violations); id. § 662 (minimum wage); id. § 680 (minimum wage—farm workers); id. § 704(8) (unfair labor practices—labor relations); id. § 736 (stress evaluators); N.Y. LAB. LAW § 741 (McKinney 2002 & Supp. 2004) (health care employees); N.Y. LAB. LAW § 880(3) (McKinney 2002) (toxic substances); N.Y. PUB. HEALTH LAW § 2801-d (McKinney 2002) (complaint of health care facility resident); id. § 2803-c (health care facilities—patients’ statement of rights); id. § 2803-d (health care facilities—reporting abuse and neglect); N.Y. SOC. SERV. LAW § 460-d(2)(vi) (McKinney 2003 & Supp. 2004) (residential care facilities—adoption of regulations); N.Y.
Employee Whistleblowing After Sarbanes–Oxley


**NORTH CAROLINA:** North Carolina recognizes a public policy exception to the at-will employment rule. See Amos v. Oakdale Knitting Co., 416 S.E.2d 166, 167 (N.C. 1992); Lenzer v. Flaherty, 418 S.E.2d 276, 287 (N.C. Ct. App. 1992). State employees are protected if they report a violation of law, fraud, misappropriation, a “[s]ubstantial and specific danger to the public health and safety,” or “[g]ross mismanagement, a gross waste of monies, or gross abuse of authority.” N.C. GEN. STAT. § 126-84 (2003); see id. §§ 126-84 to 126-88 (improper government activities); see also id. § 126-17 (equal opportunity violations—state personnel).


**NORTH DAKOTA:** North Dakota recognizes a public policy exception to the at-will employment rule. See Ressler v. Humane Soc’y, 480 N.W.2d 429, 431–32 (N.D. 1992). Private employees are protected if they report a violation of law. N.D. CENT. CODE § 34-01-20(1)(a) (Supp. 2003). Public employees are protected if they report a job-related violation of law or misuse of public resources. N.D. CENT. CODE


**OHIO:** Ohio recognizes a public policy exception to the at-will employment rule. See Pytlinski v. Borcar Prods., Inc., 760 N.E.2d 385, 387–88 (Ohio 2002). Public and private employees are protected if they report a violation of law when “the employee reasonably believes that the violation either is a criminal offense that is likely to cause an imminent risk of physical harm to persons or a hazard to public health or safety or is a felony . . . .” OHIO REV. CODE ANN. § 4113.52 (West 2001 & Supp. 2003); see id. §§ 4113.51–.53. In addition, state employees are protected if they report a violation of law or misuse of public resources. OHIO REV. CODE ANN. § 124.341 (West 2002); see also OHIO REV. CODE ANN. § 4117.11 (West 2001) (unfair labor practice); OHIO REV. CODE ANN. § 4167.13 (West 2001) (public employment risk-reduction program).

Employee Whistleblowing After Sarbanes–Oxley

2003) (retaliation against chiropractors); OHIO REV. CODE ANN. § 5101.61 (West 2001) (abuse, neglect, or exploitation of adult); OHIO REV. CODE ANN. § 5104.10 (West 2001) (child care facility employees); OHIO REV. CODE ANN. §§ 5123.604, 5123.61 (West 2001) (abuse and neglect of persons with mental retardation and developmental disabilities).


OREGON: Oregon recognizes a public policy exception to the at-will employment rule. See Dunwoody v. Handskill Corp., 60 P.3d 1135, 1142 (Or. Ct. App. 2003). It is an unlawful employment practice for an employer to retaliate against an employee for certain conduct, such as reporting “criminal activity by any person . . . .” OR. REV. STAT.
§ 659A.230 (2001). Public employees are protected if they report a violation of law, “[m]ismanagement, gross waste of funds or abuse of authority or substantial and specific danger to public health and safety resulting from action of the state, agency or political subdivision,” or “that a person receiving services, benefits or assistance from the state or agency or subdivision, is subject to a felony or misdemeanor warrant for arrest . . . .” Id. § 659A.203; see Whistleblower Law, id. §§ 659A.200–.224; see also OR. REV. STAT. § 243.672 (1999) (unfair labor practices—public employee collective bargaining).

**Area-specific whistleblower statutes:** OR. REV. STAT. § 185.170 (1999) (testimony by members of Oregon Disabilities Commission); OR. REV. STAT. § 430.755 (1999) (adult abuse); OR. REV. STAT. § 441.057 (1999) (health care facilities); id. § 441.127 (long-term care facilities); id. § 441.655 (resident abuse—long-term care facilities); OR. REV. STAT. § 443.765 (1999) (adult foster homes); OR. REV. STAT. § 464.280(3) (1999) (violations of law—gambling operations); OR. REV. STAT. § 652.220 (1999) (wage discrimination); id. § 652.355; OR. REV. STAT. § 653.060 (1999) (minimum wage); OR. REV. STAT. § 654.062 (1999) (occupational safety and health); OR. REV. STAT. § 657.260 (1999) (testimony—unemployment compensation); OR. REV. STAT. § 658.452 (1999) (farm labor employees); id. § 658.760 (farmworker camps); OR. REV. STAT. § 659A.030 (2001) (unlawful employment practices—discrimination); id. § 659A.233 (reporting certain violations; testifying at unemployment compensation hearing); id. § 659A.236 (legislative testimony); id. § 659A.259 (employee housing); OR. REV. STAT. § 663.120 (1999) (unfair labor practices—labor relations).


**Area-specific whistleblower statutes:** PA. STAT. ANN. tit. 23, § 6311 (West 2001) (child abuse); PA. STAT. ANN. tit. 24, § 5004 (West 1992 & Supp. 2003) (unfair educational practices—Pennsylvania Fair Educational Opportunities Act); PA. STAT. ANN. tit. 27, § 4112 (West Supp. 2004) (environmental laboratory employees); PA. STAT. ANN. tit. 35, § 6020.1112 (West 2003) (Hazardous Sites Cleanup Act); id. § 7130.509 (low-level waste facilities); id. § 7313 (Worker and
Employee Whistleblowing After Sarbanes–Oxley


**Area-specific whistleblower statutes:** R.I. GEN. LAWS § 5-29-15 (1999) (unprofessional conduct, incompetence, or negligence of a podiatrist); id. § 5-31.1-9 (unprofessional conduct or negligence of dentists and dental hygienists); id. § 5-37.1-5 (unprofessional conduct or negligence of physicians); R.I. GEN. LAWS § 23-1.1-14 (2001) (occupational health); id. §§ 23-17.8-4 to -5 (abuse in health care facilities); id. § 23-17.14-29 (Hospital Conversions Act); id. § 23-24.5-.11 (asbestos abatement); R.I. GEN. LAWS § 27-18-45 (2002) (retaliation by insurance companies against medical services providers); id. § 27-19-37 (retaliation by nonprofit hospital service corporations against medical services providers); id. § 27-20-32 (retaliation by nonprofit medical
service corporations against medical services providers); id. § 27-41-46 (retaliation by HMOs against medical services providers); id. § 27-54-7 (Insurance Fraud Prevention Act); id. § 27-63-1 (health care fraud); R.I. GEN. LAWS § 28-5-7 (2003) (unlawful employment practices—fair employment); id. § 28-6-21 (wage discrimination); id. § 28-7-13 (unfair labor practices—Labor Relations Act); id. § 28-12-16 (minimum wage); id. § 28-14-18 (payment of wages); id. § 28-20-21 (occupational safety); id. § 28-21-8 (testimony—Hazardous Substances Right-To-Know Act); R.I. GEN. LAWS § 40.1-27-6 (1997) (abuse of persons with developmental disabilities); R.I. GEN. LAWS § 42-66.7-8 (1998) (long-term care).


SOUTH DAKOTA: South Dakota recognizes a public policy exception to the at-will employment rule. See Dahl v. Combined Ins. Co., 621 N.W.2d 163, 166–67 (S.D. 2001). Public employees who are subjected to retaliation for reporting a violation of state law may file a grievance. S.D. CODIFIED LAWS § 3-6A-52 (Michie 2004); see also id. § 3-18-3.1 (unfair practices—public employees’ unions).

Area-specific whistleblower statutes: S.D. CODIFIED LAWS § 20-13-26 (Michie 1995) (human rights); S.D. CODIFIED LAWS § 27B-8-43 (Michie Supp. 2003) (abuse, neglect, or exploitation of a person with a developmental disability); S.D. CODIFIED LAWS § 28-1-45.7 (Michie 1999) (Older Americans ombudsman program); S.D. CODIFIED LAWS § 60-9A-12 (Michie 2004) (unfair practices—collective bargaining); S.D. CODIFIED LAWS § 60-11-17.1 (Michie 1993) (wage complaints and proceedings); id. § 60-12-21 (wage discrimination).

TENNESSEE: Tennessee recognizes a public policy exception to the
Employee Whistleblowing After Sarbanes–Oxley


Area-specific whistleblower statutes: TEX. AGRIC. CODE ANN. § 125.013 (Vernon 1995) (Agricultural Hazard Communication Act); TEX. FAM. CODE ANN. § 261.110 (Vernon 2002) (child abuse and neglect); TEX. HEALTH & SAFETY CODE ANN. § 142.0093 (Vernon 2001) (home health, hospice, and personal assistance services); id. § 161.134 (health care facilities); id. § 242.133 (nursing homes and related institutions); id. § 242.655 (cooperation with long-term care legislative oversight committee); id. § 247.068 (assisted living facilities); TEX. HEALTH & SAFETY CODE ANN. § 252.132 (Vernon 2001 & Supp. 2004) (intermediate care facilities); TEX. HEALTH & SAFETY CODE ANN. § 502.017 (Vernon 2003) (Hazard Communication Act); TEX. HUM.
RES. CODE ANN. § 36.115 (Vernon 2001) (Medicaid fraud prevention—private actions); id. § 101.064 (long-term care facilities); TEX. LAB. CODE ANN. § 21.055 (Vernon 1996) (employment discrimination); id. §§ 411.081–.083 (use of telephone hotline to report occupational safety or health violation); id. § 451.001 (testimony and claims—workers’ compensation); TEX. LOC. GOV’T CODE ANN. § 160.006 (Vernon 1999) (county employees—grievances); TEX. OCC. CODE ANN. § 160.012 (Vernon 2004) (reports concerning physicians); id. § 301.413 (reports concerning nurses; nurse reports of patient care issues); id. § 303.009 (nursing peer review); id. §§ 505.602–.603 (social workers—incidents harmful to clients).

**Case law:** Gold v. City of College Station, 40 S.W.3d 637 (Tex. App. 2001); Austin v. HealthTrust, Inc., 967 S.W.2d 400, 403 (Tex. 1998); Winters v. Houston Chronicle Publ’g Co., 795 S.W.2d 723, 724 (Tex. 1990); Sabine Pilot Serv., Inc. v. Hauck, 687 S.W.2d 733, 734 (Tex. 1985) (creating “very narrow exception to the employment-at-will doctrine,” but requiring plaintiff “to prove by a preponderance of the evidence that his discharge was for no reason other than his refusal to perform an illegal act”).

**UTAH:** Utah recognizes a public policy exception to the at-will employment rule. See Rackley v. Fairview Care Ctrs., Inc., 23 P.3d 1022, 1026–27 (Utah 2001). Public employees are protected if they report a violation of law or waste. Utah Protection of Public Employees Act, UTAH CODE ANN. §§ 67-21-1 to -9 (2000).


**VERMONT:** Vermont recognizes a public policy exception to the at-will employment rule. See Payne v. Rozendaal, 520 A.2d 586, 588 (Vt. 1986).

**Area-specific whistleblower statutes:** VT. STAT. ANN. tit. 3, § 961 (2003) (unfair labor practices—State Employees Labor Relations Act); id. § 1026 (unfair labor practices—Judiciary Employees Labor Relations Act); VT. STAT. ANN. tit. 18, § 1427 (2002) (smoking in the workplace); VT. STAT. ANN. tit. 21, §§ 231–232 (2003) (occupational safety and
Employee Whistleblowing After Sarbanes–Oxley

health); id. § 473 (parental and family leave); id. § 491 (employment rights for reserve and national guard members); id. § 494d (Polygraph Protection Act); id. § 495 (unlawful employment practices—fair employment); id. § 1621 (State Labor Relations Act); VT. STAT. ANN. § 1726 (1987 & Supp. 2003) (unfair labor practices—Vermont Municipal Labor Relations Act); VT. STAT. ANN. tit. 33, § 4920 (2001) (child abuse); id. § 6909 (abuse of vulnerable adults); id. § 7508 (long-term care).

VIRGINIA: Virginia recognizes a public policy exception to the at-will employment rule, but the state’s courts have thus far rejected its application to a “generalized, common-law ‘whistleblower’ retaliatory discharge claim.” Dray v. New Mkt. Poultry Prods., Inc., 518 S.E.2d 312, 313 (Va. 1999).

Area-specific whistleblower statutes: VA. CODE ANN. §§ 2.2-3000 to -3004 (Michie 2001) (state employees grievance procedure); VA. CODE ANN. § 8.01-581.17 (Michie Supp. 2003) (report of patient safety data to patient safety organization); VA. CODE ANN. § 15.2-1507 (Michie 2003) (local government employees—grievance procedure); VA. CODE ANN. § 32.1-125.4 (Michie 2001 & Supp. 2003) (hospitals); VA. CODE ANN. § 32.1-138.4 (Michie 2001) (nursing facilities); VA. CODE ANN. § 40.1-51.2 (Michie 2002) (participation in safety and health inspection); id. § 40.1-51.2:1 (occupational safety and health); VA. CODE ANN. § 51.5-39.10 (Michie 2002) (complaints and investigations concerning persons with disabilities); VA. CODE ANN. § 54.1-515 (Michie 2002) (asbestos, lead, and home inspection contractors and workers); VA. CODE ANN. § 63.2-1730 (Michie 2002) (assisted living and adult day care facilities; child welfare agencies); id. § 63.2-1731 (abuse and neglect at assisted living facilities, adult day-care centers, or child welfare agencies); VA. CODE ANN. § 65.2-308 (Michie 2002) (testimony and current or future claims—workers’ compensation).

Area-specific whistleblower statutes: WASH. REV. CODE § 18.20.185 (2004) (boarding homes); id. § 18.51.220 (nursing homes); id. § 18.88A.230 (nursing assistants, community based care settings, and in-home service agencies); WASH. REV. CODE § 19.30.190 (2004) (employees of farm labor contractors and agricultural employers); WASH. REV. CODE § 28B.52.073 (2004) (unfair labor practices—community colleges); WASH. REV. CODE § 41.56.140 (2004) (public employees filing unfair labor practice charge); id. § 41.59.140 (unfair labor practices—Educational Employment Relations Act); id. § 41.76.050 (unfair labor practices—higher education faculty); WASH. REV. CODE § 43.06A.085 (2004) (family and children’s ombudsman); id. § 43.70.075 (quality of care by health care provider); id. § 43.190.090 (long-term care facilities); WASH. REV. CODE § 47.64.130 (2004) (unfair labor practices—ferry system); WASH. REV. CODE § 49.12.130 (2004) (testimony—primarily concerning wages and working conditions); id. § 49.12.287 (sick leave; care of family members); id. § 49.17.160 (industrial safety and health); id. § 49.26.150 (asbestos); id. § 49.46.100 (minimum wage); id. § 49.60.210 (unfair practices—discrimination); id. § 49.78.130 (family leave); WASH. REV. CODE § 70.124.100 (2004) (nursing homes and state hospitals); WASH. REV. CODE § 73.16.032 (2004) (employment of veterans); WASH. REV. CODE § 74.34.180 (2004) (abuse of vulnerable adults); id. § 74.39A.060 (long-term care facilities).


Area-specific whistleblower statutes: W. VA. CODE ANN. § 5-11-9 (Michie 2002) (unlawful discriminatory practices); W. VA. CODE ANN. § 9-6-12 (Michie 2003) (nursing homes); W. VA. CODE ANN. § 16-5C-8 (Michie 2001) (nursing homes); W. VA. CODE ANN. § 16-5D-8 (Michie 2001 & Supp. 2004) (assisted living residences); W. VA. CODE ANN. § 16-5L-18 (Michie 2001) (long-term care); id. § 16-5N-8 (residential care communities); id. § 16-32-14 (asbestos abatement); id. § 16-39-4 (health care workers); W. VA. CODE ANN. § 21-1A-4 (Michie 2002) (unfair labor practices—private sector labor relations); id. § 21-3A-13 (occupational safety and health); id. § 21-5B-3 (wage discrimination); id. § 21-5C-7 (wages and hours); id. § 21-5E-3 (wage discrimination—state employees); W. VA. CODE ANN. § 22-3-31 (Michie 2002) (Surface Coal
Employer Whistleblowing After Sarbanes–Oxley

Mining and Reclamation Act); id. § 22A-1-22 (miners and authorized representatives).


§ 289.42 (West 2004) (solid and hazardous waste facilities).

**Wyoming:** Wyoming recognizes a narrow public policy exception to the at-will employment rule. See McLean v. Hyland Enters., 34 P.3d 1262, 1268–69 (Wyo. 2001). State employees are protected if they report a violation of law, fraud, waste, gross mismanagement, or a danger to health or safety. WYO. STAT. ANN. § 9-11-103 (Michie 2003).

Employee Whistleblowing After Sarbanes–Oxley

APPENDIX B

This Appendix summarizes the federal laws—other than the Sarbanes–Oxley Act—that either directly address whistleblowing or relate to employee disclosure of information that furthers the public welfare. The statutes are broken down by categories, such as environmental protection, that best summarize the underlying policy the whistleblowing provision is thought to protect.

As suggested by this list of statutes, the current federal regulatory approach is piecemeal. If an employee blows the whistle on a specific type of violation, then that employee is protected from retaliation by federal law. If an employee blows the whistle on a violation of a statute that is not listed below, then that employee has no federal protection and, depending on the vagaries of state law, may have no remedy whatsoever. The National Whistleblower Center, a non-profit group that advocates on behalf of whistleblowers, has pointed out this weakness in the law and is lobbying to have a broad federal protection for whistleblowers enacted. The text of the Center’s model act can be found at http://www.whistleblowers.org/html/model_whistleblower_law.html (last visited Oct. 7, 2004). It is worth noting that such a model act would be also be beneficial on a state level to provide employees uniform protections.

**Banks and Banking:** 12 U.S.C.A. § 1441a(q) (West 2001) (employees of RTC, Thrift Depositor Protection Oversight Board, and RTC contractors); 12 U.S.C. § 1790b (2000) (federal credit unions); id. § 1831j (depository institution employees).

**Employment:** 29 U.S.C. § 158(a)(4) (2000) (unfair labor practices—National Labor Relations); id. § 215(a)(3) (fair labor standards); id. § 623(d) (age discrimination); id. § 660(c) (occupational safety and health); id. § 1140 (Employee Retirement Income Security and Welfare and Pension Plans Disclosure Acts); id. § 1855(a) (migrant and seasonal agricultural workers); id. § 2002 (lie detector tests); id. § 2615(b) (family and medical leave); 30 U.S.C. § 815(c) (2000) (mine safety and health); 33 U.S.C. § 948a (2000) (testimony—longshore and harbor workers’ compensation); 42 U.S.C. § 2000e-3(a) (2000) (civil rights—equal employment opportunities); id. § 12203(a) (equal opportunity for individuals with disabilities); 46 U.S.C. § 2114(a) (2000) (seamen).


Health Care: 42 U.S.C. § 1395dd(i) (2000) (hospital employees); id. § 1997d (Civil Rights of Institutionalized Persons Act); id. § 3058g(j) (state long-term care ombudsman program—eligibility for funding).


Safety/Transportation: 46 U.S.C. app. § 1506(a) (2000) (shipping); 49 U.S.C. § 20109(a) (2000) (railroad carrier employees); id. § 31105(a) (commercial motor vehicle safety); id. § 42121(a) (airline employees);
Employee Whistleblowing After Sarbanes–Oxley


DOES DELAWARE’S SECTION 102(b)(7) PROTECT RECKLESS DIRECTORS FROM PERSONAL LIABILITY? ONLY IF DELAWARE COURTS ACT IN GOOD FAITH

Matthew R. Berry

Abstract: Section 102(b)(7) of the Delaware Corporate Code allows a corporation to amend its certificate of incorporation to exculpate directors from all duty of due care violations. The Delaware General Assembly enacted this law in response to the shrinking pool of qualified directors, which was caused by the Delaware State Supreme Court’s decision in Smith v. Van Gorkom that imposed personal liability on directors for gross negligence. Delaware courts have unequivocally stated that section 102(b)(7) protects directors against personal liability arising from gross negligence, but not against liability arising from a lack of good faith. However, Delaware courts have not provided clear guidance as to whether the statute protects directors from personal liability arising from recklessness. If Delaware courts classify reckless conduct as a breach of the duty of due care, then section 102(b)(7) protects directors from liability arising from recklessness. Conversely, if Delaware courts classify reckless conduct as a breach of the duty of good faith, then section 102(b)(7) offers reckless directors no protection. This Comment proposes that section 102(b)(7) protects directors from personal liability arising from reckless decisions for two reasons. First, recklessness is merely a subset of gross negligence in Delaware corporate law. Because section 102(b)(7) unambiguously protects directors from liability arising from gross negligence, it also protects them from liability arising from recklessness. Second, recklessness by definition is conduct that involves no intention to cause harm. Because the Delaware State Supreme Court requires an illicit motive or bad faith state of mind to establish bad faith conduct, a reckless director breaches only the duty of due care and is protected by section 102(b)(7).

Directors play a crucial role in the corporate decisionmaking process.1 They oversee and advise managerial decisions and provide a necessary check on top managerial power.2 In fulfilling these tasks, the directors owe fiduciary duties to the corporation and its shareholders,3 and face personal liability for damages arising from the breach of any of these fiduciary duties.4 However, directors in a typical corporation face time and budget constraints and cannot always make decisions based on full

2. See id.
3. See Smith v. Van Gorkom, 488 A.2d 858, 872 (Del. 1985) (stating that directors are charged with a fiduciary duty to the corporation and its shareholders when carrying out their managerial roles).
4. See Emerald Partners v. Berlin, 787 A.2d 85, 90 (Del. 2001) (stating that directors are personally liable for monetary damages if they breach any of their fiduciary duties).
and accurate information. Holding directors personally liable for negligence, or even gross negligence, in the decisionmaking process might harm corporations—many qualified directors may become reluctant to serve on boards because they desire to limit their personal exposure to liability. Those directors who do serve on boards would become increasingly risk-averse causing them to avoid risky projects with potentially high returns for the corporation.

The Delaware General Assembly enacted section 102(b)(7) of the Delaware Corporate Code to provide directors with additional protection from personal liability. The legislators attempted to remedy directors’ difficult position of making material decisions on incomplete information, and to encourage them to act in the shareholders’ best interests. Section 102(b)(7) provides shareholders with an option to amend their corporation’s certificate of incorporation to include an exculpation provision. Delaware courts have interpreted this provision to protect directors from personal liability arising from simple and gross negligence. Further, section 102(b)(7) provides that directors are not

6. See id. at 449.
8. See Veasey, supra note 7, at 693–94.
9. DEL. CODE ANN. tit. 8, § 102(b)(7) (2003). Section 102(b) reads in relevant part:
In addition to the matters required to be set forth in the certificate of incorporation by subsection (a) of this section, the certificate of incorporation may also contain any or all of the following matters:

7. A provision eliminating or limiting the personal liability of a director to the corporation or its stockholders for monetary damages for breach of fiduciary duty as a director, provided that such provision shall not eliminate or limit the liability of a director:
(i) For any breach of the director’s duty of loyalty to the corporation or its stockholders;
(ii) for acts or omissions not in good faith or which involve intentional misconduct or a knowing violation of law; (iii) under § 174 of this title; or (iv) for any transaction from which the director derived an improper personal benefit.

Id.
11. DEL. CODE ANN. tit. 8, § 102(b)(7); Veasey et al., supra note 1, at 402.
Personal Liability of Reckless Directors

protected from liability arising from their bad faith conduct. However, the courts have not provided clear guidance as to whether section 102(b)(7) protects directors from personal liability arising from their reckless conduct.

This Comment argues that the Delaware State Supreme Court should interpret section 102(b)(7) to exculpate directors from personal liability arising from reckless conduct because recklessness is a breach of the duty of due care. Part I discusses directors’ fiduciary duties. Part II outlines how Delaware courts have interpreted section 102(b)(7) to alter directorial liability arising from breaches of those fiduciary duties. Part III posits two arguments that section 102(b)(7) protects directors from liability arising from their reckless conduct. First, Part III.A argues that section 102(b)(7) protects directors from liability arising from recklessness because recklessness is a subcategory of gross negligence under Delaware corporate law. Next, Part III.B argues that section 102(b)(7) protects directors from liability arising from recklessness because reckless directors lack the bad faith intent that Delaware courts require to establish a breach of the duty of good faith. Finally, Part IV concludes that section 102(b)(7) protects directors from liability arising from recklessness because recklessness is a breach of the duty of due care.

I. UNDER DELAWARE COMMON LAW, DIRECTORS OWE SHAREHOLDERS DUTIES OF GOOD FAITH, LOYALTY, AND DUE CARE

Delaware common law charges directors with an “unyielding fiduciary duty to the corporation and its shareholders.” This fiduciary duty is commonly divided into the duty of good faith, duty of loyalty, and duty of due care. If directors breach any of these three fiduciary duties, then the shareholders can hold them personally liable for resulting monetary damages.

14. See Veasey et al., supra note 1, at 403.
A. Directors Owe Shareholders a Duty of Good Faith

Delaware courts do not agree whether the duty of good faith is an independent fiduciary duty or merely a part of the duty of loyalty, but the courts do agree that directors who act in bad faith are personally liable for any resulting damages. To establish bad faith, the Delaware State Supreme Court requires proof that the directors acted with a bad faith motive. However, the Delaware Court of Chancery recently has held that directors could act in bad faith without proof that the directors had a bad faith motive.

1. The Delaware State Supreme Court Requires Proof of a Bad Faith Motive Before Finding a Breach of the Duty of Good Faith

The Delaware State Supreme Court evaluates directors’ motives to distinguish breaches of the duty of due care from breaches of the duty of good faith. For example, in Zirn v. VLI Corporation, the Delaware State Supreme Court requires proof that the directors acted with a bad faith motive. Other Delaware courts have implied that the duty of good faith is an independent duty. Other Delaware courts have implied that the duty of good faith is an independent duty. See, e.g., Cede & Co., 634 A.2d at 361 (dividing directors’ fiduciary duties into the duty of good faith, the duty of due care, and the duty of loyalty); see also Hillary A. Sale, Delaware’s Good Faith, 89 CORNELL L. REV. 456, 483–84 (2004) (asserting that recent Delaware decisions have treated the duty of good faith as an independent fiduciary duty). Whether the duty of good faith is independent from the duty of loyalty is beyond the scope of this Comment.

18. Some Delaware courts have concluded that the duty of good faith is merely a subset of the duty of loyalty. See, e.g., Guttman v. Huang, 823 A.2d 492, 506 n.34 (Del. Ch. 2003) (stating that the duty of good faith is not independent of the duty of loyalty); Nagy v. Bistricer, 770 A.2d 43, 49 n.2 (Del. Ch. 2000) (“By definition, a director cannot simultaneously act in bad faith and loyally towards the corporation and its stockholders.”). Other Delaware courts have implied that the duty of good faith is an independent duty. See, e.g., Cede & Co., 634 A.2d at 361 (dividing directors’ fiduciary duties into the duty of good faith, the duty of due care, and the duty of loyalty); see also Hillary A. Sale, Delaware’s Good Faith, 89 CORNELL L. REV. 456, 483–84 (2004) (asserting that recent Delaware decisions have treated the duty of good faith as an independent fiduciary duty).

19. See, e.g., Malpiede v. Townson, 780 A.2d 1075, 1095 (Del. 2001) (stating that the purpose of section 102(b)(7) was to protect directors from duty of due care claims and not duty of good faith or duty of loyalty claims).

20. See, e.g., Zirn v. VLI Corp., 681 A.2d 1050, 1061–62 (Del. 1996) (stating that the directors acted in good faith because they lacked any motive to deceive or harm the corporation or its shareholders).

21. Delaware’s use of chancery courts is unusual in the Anglo-American judicial system. The Delaware chancery court is a court of equity consisting of one chancellor and four vice chancellors, all of whom are selected by the governor. See Cyril Moscow, Michigan or Delaware Incorporation, 42 WAYNE L. REV. 1897, 1917 (1996). The chancery court hears all equitable cases, but is especially capable of adjudicating complex corporate issues given each chancellor’s extensive experience in corporate law. Id. See generally William T. Quillen & Michael Hanrahan, A Short History of the Delaware Court of Chancery, 18 DEL. J. CORP. L. 819 (1993).

22. See, e.g., In re Walt Disney Co. Derivative Litig., 825 A.2d 275, 290 (Del. Ch. 2003) (concluding that the plaintiff could proceed on her breach of duty of good faith claim without proof of directors’ motives).

23. See, e.g., Zirn, 681 A.2d at 1061–62 (stating that the directors acted in good faith because they lacked any motive to deceive or harm the corporation or its shareholders); Smith v. Van
State Supreme Court required the plaintiff to demonstrate that directors had an illicit motive before the court would conclude that the directors acted in bad faith.25 In *Zirn*, the court considered whether directors’ material misrepresentations made during a tender offer breached the duty of good faith.26 Instead of focusing on the severity of the misrepresentation or the resulting harm to the shareholders, the court evaluated whether the directors acted with a bad faith motive.27 The court concluded that the directors acted in good faith because they lacked any motive to deceive the shareholders.28

Similarly, in *Smith v. Van Gorkom*,29 the Delaware State Supreme Court approached the question of good faith by stating that it need not inquire into the directors’ motives unless there were allegations or proof of bad faith conduct.30 In *Van Gorkom*, the court held that the directors breached their duty of due care by relying too heavily on the chief executive officer’s judgment and failing to perform an adequate independent analysis.31 In dicta, the court noted that because the plaintiff had not alleged or provided proof of bad faith conduct, the directors’ motives were irrelevant.32 This suggests that, had the plaintiff alleged that the directors acted in bad faith, a court would have to evaluate the directors’ motives.

The Delaware State Supreme Court has also evaluated defendants’ motives to define bad faith conduct outside the corporate director context.33 In *Desert Equities, Inc. v. Morgan Stanley Leveraged Equity Fund*, 488 A.2d 858, 873 (Del. 1985) (stating that because there were no allegations that the directors breached the duty of good faith, the directors’ motives were irrelevant).


25. See id. at 1061–62.

26. Id. at 1053–54.

27. See id.

28. Id. (“The record reveals that any misstatements or omissions that occurred were made in good faith. The VLI directors lacked any pecuniary motive to mislead the VLI stockholders intentionally and no other plausible motive for deceiving the stockholders has been advanced.” (emphasis added)).

29. 488 A.2d 858 (Del. 1985).

30. See id. at 873.

31. See id. at 893.

32. See id. at 873 (“[T]here were no allegations of fraud, bad faith, or self-dealing, or proof thereof. Hence, it is presumed that the directors reached their business judgment in good faith, and considerations of motive are irrelevant to the issue before us.” (emphasis added) (internal citation omitted)).

33. See, e.g., Desert Equities, Inc. v. Morgan Stanley Leveraged Equity Fund, II, L.P., 624 A.2d 1199, 1206 (Del. 1993) (evaluating a general partner’s state of mind to determine if he acted in bad
Fund, II, L.P., the court concluded that a bad faith motive is a required element of bad faith conduct. The court considered whether a general partner acted in bad faith by excluding the plaintiff from participating in its investment funds. The court reasoned that a bad faith claim depends upon the defendant’s “tortious state of mind.” The court defined bad faith as “not simply bad judgment or negligence, but rather . . . the conscious doing of a wrong because of dishonest purpose or moral obliquity.” Distinguishing bad faith from negligence, the court reasoned that bad faith “contemplates a state of mind affirmatively operating with furtive design or ill will.” Therefore, the court concluded that a “dishonest purpose” and “ill will” are required elements of a bad faith claim.

2. Despite Delaware State Supreme Court Precedent, Some Delaware Chancery Courts Have Found a Breach of the Duty of Good Faith Even Without Proof of a Bad Faith Motive

Although the Delaware State Supreme Court requires a bad faith motive to prove a breach of the duty of good faith, Delaware chancery courts have not consistently enforced this requirement. For example, the Delaware chancery court in In re Caremark International, Inc.
Personal Liability of Reckless Directors

Derivative Litigation noted in dicta that directors’ sustained or systematic failure to oversee the corporate decisionmaking process is sufficient to prove bad faith even without proof of a bad faith motive. Caremark involved directors who allegedly failed to monitor adequately the corporation’s compliance with the law. The Caremark chancellor stated that, to prevail, the plaintiffs must prove that the directors knew or should have known that illegal activities were occurring, that the directors made no good faith effort to prevent or address the situation, and that their failure to act proximately caused the damage. However, the chancellor did not require the plaintiff to prove that the directors had a bad faith motive or intent.

Consistent with Caremark, the Delaware chancery court recently concluded that directors’ conscious and sustained failure to fulfill their oversight duties was sufficient to establish bad faith in In re Walt Disney Co. Derivative Litigation. The Disney court considered whether Disney’s directors acted in bad faith by allowing the chief executive officer (CEO), Michael Eisner, to unilaterally negotiate an employment contract with Disney’s new president, Michael Ovitz. The directors had almost no involvement in either the decision to hire Ovitz or the drafting of his employment agreement. Instead, the directors granted Eisner full power to negotiate and approve the final terms and conditions of the contract without the directors’ further approval. A year after Disney hired Ovitz, both Eisner and the directors realized their mistake, but the board again did not oversee Ovitz’s termination. Without the directors’ oversight, Eisner granted Ovitz a non-fault termination agreement that enabled Ovitz to realize $100,000,000 in compensation from stock options.

---

43. 698 A.2d 959 (Del. Ch. 1996).
44. See id. at 971. Caremark is often credited for establishing the foundation for imposing personal liability on directors for a breach of the duty of good faith. See Sale, supra note 18, at 468–69.
45. Caremark, 698 A.2d at 960–62.
46. Id. at 971.
47. See id.
48. 825 A.2d 275, 278 (Del. Ch. 2003).
49. See id. at 279–81.
50. See id. at 287.
51. Id. at 281.
52. See id. at 282–85.
53. Id. at 284.
Based upon this directorial misconduct, the chancellor concluded that the plaintiffs were entitled to proceed with their claim that the Disney directors breached the duty of good faith. The chancellor found that the plaintiffs sufficiently pleaded that the directors “consciously and intentionally disregarded their responsibilities,” acted with “deliberate indifference,” and “adopt[ed] a ‘we don’t care about the risks’ attitude . . . .” The chancellor stated that when directors abdicate all responsibility in making a material corporate decision, it raises the question of whether the board’s decisionmaking process was conducted in good faith. Without citing any authority, the chancellor concluded that if directors caused harm to shareholders by consciously ignoring their corporate duties, then the directors acted in bad faith or committed intentional misconduct.

Unlike the Delaware State Supreme Court in Van Gorkom, Zirn, or Desert Equities, the Disney chancellor did not inquire into the directors’ intent or motives. Instead, the chancellor assumed that, given the severity of the directors’ conduct and the resulting harm, the directors must have acted in bad faith. The Disney chancellor also avoided addressing whether recklessness is sufficient to prove bad faith. In fact, the chancellor never used the term “reckless,” even though the plaintiffs alleged that the board members “consciously and intentionally disregarded their responsibilities, adopting a ‘we don’t care about the risks’ attitude concerning a material corporate decision.”

In sum, although courts agree that directors who act in bad faith are liable for any resulting damages, the Delaware State Supreme Court

54. See id. at 278.
55. Id. at 289 (emphasis omitted).
56. Id.
57. Id.
58. Id. at 289–90.
59. Id. at 290 (“Where a director consciously ignores his or her duties to the corporation, thereby causing economic injury to its stockholders, the director’s actions are either ‘not in good faith’ or ‘involve intentional misconduct.’” (quoting DEL. CODE ANN. tit. 8, § 102(b)(7)(ii) (2003))).
61. See Disney, 825 A.2d at 289.
62. Id. (emphasis omitted).
63. See, e.g., Malpiede v. Townson, 780 A.2d 1075, 1094–95 (Del. 2001) (stating that the purpose of section 102(b)(7) was to protect directors from duty of care claims and not from duty of good faith or duty of loyalty claims).
Personal Liability of Reckless Directors

has not clearly defined the duty of good faith. The court expressly requires a bad faith motive to establish bad faith. However, some recent chancery court decisions have concluded that directors acted in bad faith even without proof of a bad faith motive.

B. Directors Owe Shareholders a Duty of Loyalty

Although the duty of good faith still floats in murky waters, Delaware courts have consistently described the duty of loyalty as the conflict between the directors’ duty to the corporation and their self-interest. Within the duty of loyalty is an affirmative duty to protect the corporation’s interests and avoid any conduct that would injure the corporation or its shareholders or deprive either of a benefit. Directors breach the duty of loyalty when their economic or other interests conflict with the corporation or its shareholders’ interests.

C. Directors Owe Shareholders a Duty of Due Care

The duty of due care requires that directors exercise an informed business judgment when making corporate decisions and authorizes courts to impose personal liability on directors for negligence or recklessness. If directors make a well-informed business decision, courts apply the business judgment rule, which limits judicial review to scrutiny of directors’ decisionmaking process rather than to the substantive decision itself. Courts interpret the business judgment rule to protect directors from liability arising from negligence, but not from gross negligence. Because Delaware courts consider recklessness to be

---

64. See, e.g., Sale, supra note 18, at 468 (“[E]xplication of the good faith obligations—indeed, even the use of that term—is dicta.”).
65. See, e.g., Zirn, 681 A.2d at 1061–62 (stating that the directors acted in good faith because they lacked any motive to deceive or harm the corporation or its shareholders).
66. See, e.g., Disney, 825 A.2d at 290 (holding that the plaintiff could proceed on her breach of the duty of good faith claim without proof of the directors’ motives).
68. Id.; see also 1 RODMAN WARD, JR., EDWARD P. WELCH & ANDREW J. TUREZYN, FOLK ON THE DELAWARE GENERAL CORPORATION LAW § 141.2.1.1 (4th ed. 2004) (describing the duty of loyalty).
69. WARD ET AL., supra note 68, § 141.2.1.1.
72. Van Gorkom, 488 A.2d at 873.
a subset of gross negligence, the business judgment rule similarly does not protect directors from liability arising from recklessness.


The business judgment rule offers directors limited protection from personal liability arising from business decisions. The rule creates a presumption that in making a business decision, directors acted on an informed basis, in good faith, and with the honest belief that their decision was in the corporation’s best interest. In light of this presumption, courts may evaluate only the decisionmaking process and not the substantive decision itself. However, courts apply the business judgment rule only when the directors actually make a decision. Thus, when directors abdicate their duties or fail to make a conscious decision, the rule provides no protection.

2. The Business Judgment Rule Protects Directors from Liability Arising from Negligence, but Not Gross Negligence

The business judgment rule protects directors from liability arising from negligence, but not from gross negligence. This limitation of the business judgment rule became clear when the Delaware State Supreme Court imposed personal liability on directors for their grossly negligent conduct in Van Gorkom. The directors in Van Gorkom approved a proposed sale of the company without prior consideration of the sale,

---

74. See, e.g., Cede & Co., 634 A.2d at 361 (noting that a plaintiff must rebut the business judgment rule before directors can be held personally liable).
75. See Van Gorkom, 488 A.2d at 872.
76. See Cede & Co., 634 A.2d at 361.
78. See id.
79. Van Gorkom, 488 A.2d at 873.
80. See id. at 893.
Personal Liability of Reckless Directors

even though there were no time constraints, the directors approved amendments to the merger agreement without reading them; and the directors gave the CEO authority to execute the merger without further board approval. Moreover, when the CEO did execute the merger, he did so at a party without reading the final agreement. Because the directors’ decision did not involve fraud, bad faith, or self-dealing, the court analyzed the decision under the business judgment rule. However, the business judgment rule protects directors only against personal liability arising from negligence, and the court held that the directors’ conduct amounted to gross negligence. Accordingly, the court imposed personal liability on the directors for their grossly negligent conduct.

3. The Business Judgment Rule Does Not Protect Directors from Liability Arising from Recklessness Because Recklessness Is a Subset of Gross Negligence

The business judgment rule does not protect directors from liability arising from reckless conduct because Delaware courts deem corporate recklessness as a subset of gross negligence. “Recklessness” is the “knowing disregard of a substantial and unjustifiable risk.” Reckless actors have no intent to cause harm; instead they have an “I don’t care” attitude. While most courts distinguish recklessness from gross negligence, Delaware courts depart from the majority view by

81. See id. at 874.
82. Id. at 869.
83. See id. at 879.
84. Id.
85. See id. at 873.
86. See id. at 873, 893 (stating that the relevant standard is gross negligence and that the directors breached that standard).
87. See id. at 893.
91. See Del. P.J.I. Civ. § 5.9; see also Dooley, 2003 WL 1903771, at *6 (applying the definition of recklessness from Del. P.J.I. Civ. § 5.9).
92. See Allen et al., supra note 5, at 453 (noting that the gross negligence standard in Delaware
designating corporate recklessness as a subset of gross negligence. 93 For example, in Tomczak v. Morton Thiokol, Inc., 94 a Delaware chancellor considered whether directors breached their fiduciary duties in approving a sale of one of the corporation’s divisions. 95 Upon holding that the directors acted in good faith, the chancellor stated, “In the corporate context, gross negligence means ‘reckless indifference to or a deliberate disregard of the whole body of stockholders’ or actions which are ‘without the bounds of reason.’” 96 Other Delaware courts have also concluded that gross negligence includes corporate recklessness. 97 In addition to Delaware case law, three Delaware chancellors 98 concluded in a recent law review article that plaintiffs must establish that directors acted recklessly to prove that they acted with gross negligence. 99 Moreover, one of the leading Delaware corporate law treatises agreed with the three chancellors that recklessness is merely an element of a gross negligence claim in Delaware corporate jurisprudence. 100 Because the business judgment rule offers directors no protection from liability arising from gross negligence, 101 it similarly offers them no protection

corporate cases is harder to establish than the gross negligence standard courts normally apply in tort or criminal cases); see also Rabkin v. Philip A. Hunt Chem. Corp., 547 A.2d 963, 970 (Del. Ch. 1986) (stating that Delaware corporate law has a different standard for gross negligence than other areas of the law).

93. See, e.g., Tomczak, 1990 WL 42607, at *12 (“In the corporate context, gross negligence means ‘reckless indifference to or a deliberate disregard of the whole body of stockholders’ or actions which are ‘without the bounds of reason.’”) (quoting Allaun v. Consol. Oil Co., 147 A. 257, 261 (Del. Ch. 1929) (emphasis added)); see also Rabkin, 547 A.2d at 970 (noting that corporate gross negligence is similar to recklessness); Allen et al., supra note 5, at 453 (noting that courts in Delaware corporate cases have adopted a gross negligence standard that is similar to the recklessness standard).


95. Id. at *12.

96. Id. (quoting Allaun, 147 A. at 261 (emphasis added)).

97. See, e.g., Rabkin, 547 A.2d at 970 (noting that corporate gross negligence is similar to recklessness); see also WARD ET AL., supra note 68, § 141.2.2.4 (noting that corporate gross negligence is similar to recklessness); Allen et al., supra note 5, at 453 (same).

98. William T. Allen is a former vice-chancellor of the Delaware Court of Chancery. Jack B. Jacobs is a former vice-chancellor of the Delaware Court of Chancery and is currently a justice on the Delaware State Supreme Court. Leo E. Strine, Jr. is currently a vice-chancellor on the Delaware Court of Chancery.

99. See Allen, supra note 5, at 453 (“Delaware corporate cases have adopted a gross negligence standard that requires a plaintiff to demonstrate a degree of culpability on the part of the directors that is akin to the recklessness standard employed in other contexts.”).

100. See WARD ET AL., supra note 68, § 141.2.2.4.

Personal Liability of Reckless Directors

from liability arising from reckless conduct.102

II. SECTION 102(b)(7) PROTECTS DIRECTORS FROM LIABILITY ARISING FROM BREACHES OF ONLY THE DUTY OF DUE CARE

The Delaware legislature enacted section 102(b)(7)103 in response to Van Gorkom to allow shareholders to further protect directors from personal liability.104 Although the statute’s protections are broad in that it protects directors from liability arising from all breaches of the duty of due care,105 it expressly states that it offers directors no protection if they do not act in good faith.106 However, courts have not clarified whether section 102(b)(7) also protects directors from liability arising from recklessness.107

A. The Delaware Legislature Enacted Section 102(b)(7) in Response to Van Gorkom

After Van Gorkom, the Delaware legislature responded to directors’ fears of potential liability and shareholders’ desire to retain quality, risk-taking directors by enacting section 102(b)(7).108 The Delaware State Supreme Court’s decision in Van Gorkom, which held directors personally liable for shortcomings in the merger approval process, sent shock waves of apprehension throughout the corporate world.109 This apprehension led many qualified directors to resign out of fear of future personal liability, which caused the pool of qualified directors to decrease substantially.110 Even when a corporation could find qualified directors, the directors knew they might be personally liable if a court declared their conduct grossly negligent.111 This exposure to liability caused many directors to become risk-averse in corporate

102. See id.
104. See Veasey, supra note 7, at 693.
107. See infra Part II.C.
108. See Veasey, supra note 7, at 693.
109. See Veasey, supra note 7, at 692.
110. See Veasey et al., supra note 1, at 401.
111. See Allen et al., supra note 5, at 449–51.
decisionmaking.112 While investors want directors who are diligent and honest,113 they also want directors who are bold and willing to take business risks.114 To address these shareholder concerns, the Delaware legislature enacted section 102(b)(7).115

B. Section 102(b)(7) Protects Directors from Liability Arising from Breaches of the Duty of Due Care, but Not from Breaches of the Duty of Good Faith

To determine whether section 102(b)(7) protects directors from liability arising from alleged misconduct, courts must discern whether the misconduct amounts to a breach of the duty of due care or a breach of the duty of good faith.116 Section 102(b)(7) expressly states that it offers directors no protection from personal liability arising from acts not in good faith.117 However, Delaware courts interpret section 102(b)(7) to protect directors from personal liability arising from all duty of due care violations.118 Directors breach their duty of due care through negligence or gross negligence.119 The business judgment rule protects directors from liability arising from negligence,120 and section 102(b)(7) extends this protection to liability arising from gross negligence.121 Therefore, if a director can characterize a plaintiff’s claim as alleging only a breach of the duty of due care, then the director can invoke section 102(b)(7)122 to have the complaint dismissed.123

112. See Veasey, supra note 7, at 693–94.
113. Veasey, supra note 7, at 694.
114. See Veasey, supra note 7, at 684–85 (“Requiring perfect, fail-safe systems can be far more costly than any potential business loss to the stockholders.”).
115. See Veasey, supra note 7, at 693.
118. See Malpiede v. Townson, 780 A.2d 1075, 1093 (Del. 2001); Zirn v. VLI Corp., 681 A.2d 1050, 1061 (Del. 1996).
119. See Aronson v. Lewis, 473 A.2d 805, 812 (Del. 1984) (stating that directors receive no business judgment rule protection if they fail to make a conscious decision).
121. See Malpiede, 780 A.2d at 1094–95 (concluding that section 102(b)(7) protects directors from gross negligence).
122. This assumes that the corporation had amended its certificate of incorporation to include section 102(b)(7).
123. Malpiede, 780 A.2d at 1094–95.
Personal Liability of Reckless Directors

C. Courts Are Divided as to Whether Section 102(b)(7) Protects Directors from Personal Liability Arising from Reckless Conduct

Delaware courts agree that section 102(b)(7) protects directors from liability arising from breaches of the duty of due care, and the statute explicitly states that it offers no protection from liability arising from breaches of the duty of good faith. However, courts have not reached a definitive conclusion as to whether the statute protects directors from liability arising from recklessness. If courts conclude that recklessness is a breach of the duty of due care, then section 102(b)(7) will protect directors from liability arising from recklessness. Conversely, if courts conclude that recklessness is a breach of the duty of good faith, then section 102(b)(7) will offer directors no protection. Thus, the issue is whether recklessness is a breach of the duty of good faith or a breach of the duty of due care.

1. Delaware Courts Provide Little Guidance on Applying Section 102(b)(7) to Recklessness

The Delaware State Supreme Court has suggested that section 102(b)(7) will protect directors from liability arising from recklessness. For example, in Van Gorkom, the court implicitly determined that reckless conduct does not breach the duty of good faith. The court characterized the Van Gorkom directors’ conduct as reckless, but concluded that there was no evidence that the directors had breached their duty of good faith. Although decided before the enactment of section 102(b)(7), the Van Gorkom court’s analysis

124. See, e.g., id. at 1093 (concluding that section 102(b)(7) protects directors from gross negligence).
126. See Veasey et al., supra note 1, at 403.
127. See Malpiede, 780 A.2d at 1093 (observing that section 102(b)(7) protects directors from personal liability arising from duty of due care claims).
128. See DEL. CODE ANN. tit. 8, § 102(b)(7).
129. See, e.g., Smith v. Van Gorkom, 488 A.2d 858, 871, 873 (Del. 1985) (describing the directors’ conduct as reckless, but concluding that the conduct breached only the duty of due care, not the duty of good faith).
130. See id. at 873.
131. Id. at 871 (stating that the lower court erred by finding that the directors had not acted “recklessly or improvidently”).
132. Id. at 873.
indicates that directors’ reckless conduct does not breach their duty of good faith.

Contrary to the Van Gorkom court’s decision, the Disney chancery court suggested that directors’ reckless conduct breaches the duty of good faith.\textsuperscript{133} The Disney chancellor avoided the term “reckless” throughout the opinion, but described the directors’ conduct with terms synonymous with recklessness.\textsuperscript{134} For example, the chancellor stated that the directors “consciously and intentionally disregarded their responsibilities” and “adopt[ed] a ‘we don’t care about the risks attitude.’”\textsuperscript{135} The chancellor concluded that section 102(b)(7) provided the directors no protection from liability because their actions breached the duty of good faith.\textsuperscript{136}

However, only two years before the Disney decision, the Delaware Court of Chancery in Emerald Partners v. Berlin\textsuperscript{137} expressly questioned the argument that recklessness breaches the duty of good faith.\textsuperscript{138} In Emerald Partners, the plaintiff alleged that the directors recklessly approved a merger.\textsuperscript{139} The plaintiff argued that section 102(b)(7) should not protect the directors from liability because the directors’ reckless indifference amounted to a breach of the duty of good faith.\textsuperscript{140} However, the chancellor declined to accept this argument and responded that, at best, the plaintiff cited equivocal authority to prove that reckless indifference breaches the duty of good faith.\textsuperscript{141} The chancellor ultimately concluded that he need not determine whether reckless indifference breaches the duty of good faith because the plaintiff failed to prove that the directors acted recklessly.\textsuperscript{142}

\begin{footnotesize}
\begin{enumerate}
\item See In re Walt Disney Co. Derivative Litig., 825 A.2d 275, 278, 290 (Del. Ch. 2003).
\item See id. at 289.
\item Id.
\item See id. at 289–90.
\item See id. at *26.
\item See id. at *26 n.66.
\item Id. at **17, 18.
\item See id. at *26 n.66.
\item See id. at *26.
\item Id. at *26 n.66.
\end{enumerate}
\end{footnotesize}
Personal Liability of Reckless Directors

2. Federal Appellate Courts Have Held that Section 102(b)(7) Does Not Protect Directors from Liability Arising from Reckless Conduct

Given Delaware’s significant influence on corporate law throughout the United States, other jurisdictions have considered recklessness and good faith when interpreting section 102(b)(7) or its local equivalent. For example, in McCall v. Scott, the United States Court of Appeals for the Sixth Circuit held that section 102(b)(7) did not protect the McCall directors from personal liability arising from reckless conduct. The McCall directors allegedly knew that the senior managers’ policies encouraged employees to engage in illegal activities, but the directors recklessly failed to address these policies. The court found that the directors’ “sustained inattention to their management obligations” and “intentional ignorance of” and “willful blindness” to “red flags” surpassed mere breaches of the duty of due care. The Sixth Circuit concluded that “[u]nder Delaware law, the duty of good faith may be breached where a director consciously disregards his duties to the corporation, thereby causing its stockholders to suffer.”

The Sixth Circuit’s reference to Delaware law on the conscious disregard of directorial duties came entirely from a footnote in Nagy v. Bistrizer. In Nagy, the Delaware chancery court discussed whether directors breached their fiduciary duties by merging the corporation with

143. See generally William H. Rehnquist, The Prominence of the Delaware Court of Chancery in the State-Federal Joint Venture of Providing Justice, 48 BUS. LAW. 351 (1992) (stating that the Delaware Court of Chancery plays an important part in the United States legal system and that most United States corporations have formed in Delaware because the court has the highest standards of responsibility and excellency).
144. See, e.g., McCall v. Scott, 239 F.3d 808, 1000–01 (6th Cir. 2001) [hereinafter McCall I] (adjudicating corporate liability under section 102(b)(7) for a corporation that had amended its certificate of incorporation to include section 102(b)(7)), amended by 250 F.3d 997 (6th Cir. 2001) [hereinafter McCall II].
145. 239 F.3d 808 (6th Cir. 2001), amended by 250 F.3d 997 (6th Cir. 2001).
146. The defendant was incorporated in Delaware and had amended its certificate of incorporation to include section 102(b)(7), which required that the Sixth Circuit apply Delaware law. See id. at 1000.
147. McCall II, 250 F.3d at 1000–01.
149. McCall II, 250 F.3d at 1001 (internal quotations omitted).
150. Id. (citation omitted).
151. See id.
152. 770 A.2d 43 (Del. Ch. 2000).
a second corporation controlled by the directors.\textsuperscript{153} While listing the plaintiff's allegations in the factual background section,\textsuperscript{154} the chancellor discussed the relationship between the duty of loyalty and the duty of good faith in a footnote.\textsuperscript{155} In the footnote, the chancellor concluded that the duty of good faith is subsumed within the duty of loyalty,\textsuperscript{156} and, at best, it is a reminder that "regardless of his motive, a director who consciously disregards his duties to the corporation and its stockholders may suffer a personal judgment for monetary damages for any harm he causes."\textsuperscript{157} The chancellor provided no authority for this statement.

Two years after the \textit{McCall} decision, the United States Court of Appeals for the Seventh Circuit, in \textit{In re Abbott Laboratories},\textsuperscript{158} followed \textit{McCall} by concluding that evidence of directors' reckless conduct is sufficient to prove bad faith.\textsuperscript{159} The plaintiffs in \textit{Abbott} accused the directors of acting grossly negligent, recklessly, and intentionally when the directors failed to address employee misconduct that led to repeated violations of FDA regulations.\textsuperscript{160} The Seventh Circuit quoted nearly every line from the \textit{McCall} decision that discussed Delaware's recklessness standard and concluded that the corporation's section 102(b)(7)-like provision\textsuperscript{161} did not protect the directors from liability arising from their recklessness.\textsuperscript{162} In fact, the Seventh Circuit relied so heavily on the Sixth Circuit's interpretation of Delaware law that it cited to the \textit{McCall} decision even when it relied on the \textit{Nagy} quote that recklessness is sufficient to prove bad faith under Delaware law.\textsuperscript{163}

In sum, the extent of section 102(b)(7) protections is still unclear.\textsuperscript{164} Courts agree that section 102(b)(7) protects directors from liability for all breaches of the duty of due care\textsuperscript{165} and that it offers no protection

\begin{footnotes}
153. \textit{Id. at} 46.
154. \textit{Id. at} 49.
155. \textit{Id. at} 49 n.2.
156. \textit{Id.}
157. \textit{Id.}
158. 325 F.3d 795 (7th Cir. 2003).
159. \textit{See id. at} 811.
160. \textit{Id.}
161. The Seventh Circuit was interpreting an Illinois statute, 805 ILL. COMP. STAT. 5/8.65, that was modeled after Delaware’s section 102(b)(7). \textit{Id. at} 810.
162. \textit{Id. at} 811.
163. \textit{Id.} (citing \textit{McCall II}, 250 F.3d 997, 1001 (6th Cir. 2001), but applying the \textit{Nagy} quote).
164. \textit{See Veasey et al., supra note} 1, at 403.
\end{footnotes}
Personal Liability of Reckless Directors

from liability for breaches of the duty of good faith.\textsuperscript{166} However, courts are divided over whether section 102(b)(7) protects directors from liability arising from recklessness.\textsuperscript{167}

III. \textbf{SECTION 102(b)(7) PROTECTS DIRECTORS FROM PERSONAL LIABILITY ARISING FROM RECKLESSNESS}

Under Delaware State Supreme Court precedent, section 102(b)(7) should protect directors from personal liability arising from recklessness for two reasons. First, Delaware corporate law includes recklessness within its gross negligence standard.\textsuperscript{168} Given that Delaware courts agree that section 102(b)(7) protects directors from liability arising from gross negligence,\textsuperscript{169} the statute should accordingly protect directors from liability arising from recklessness, which is a subset of gross negligence.\textsuperscript{170} Second, even if recklessness is not a subset of gross negligence, recklessness does not breach the duty of good faith because reckless actors lack a bad faith motive.\textsuperscript{171} Without this requisite motive, recklessness is not a breach of the duty of good faith,\textsuperscript{172} but is instead a breach of the duty of due care.\textsuperscript{173} As a breach of the duty of due care, section 102(b)(7) protects directors from liability arising from recklessness.\textsuperscript{174}

---

\textsuperscript{166} See, e.g., \textit{In re Walt Disney Co. Derivative Litig.}, 825 A.2d 275, 286 (Del. Ch. 2003) (stating that section 102(b)(7) does not protect directors from personal liability arising from conduct not undertaken in good faith).


\textsuperscript{169} See, e.g., \textit{Malpiede}, 780 A.2d at 1089–90 (stating that section 102(b)(7) protects directors from liability even if the complaint is construed to allege gross negligence).

\textsuperscript{170} See, e.g., \textit{Tomczak}, 1990 WL 42607, at *12 (stating that gross negligence includes recklessness).


\textsuperscript{172} Also, it does not implicate the duty of loyalty. See \textit{supra} Part I.B.

\textsuperscript{173} If it does not fall under the duties of loyalty or good faith, the only remaining option is the duty of due care. See, e.g., \textit{Cede & Co. v. Technicolor, Inc.}, 634 A.2d 345, 361 (Del. 1993) (discussing the three components of fiduciary duty: duty of loyalty, duty of care, and duty of good faith).

\textsuperscript{174} See, e.g., \textit{Malpiede}, 780 A.2d at 1092–93 (noting that section 102(b)(7) protects directors from liability arising from all breaches of the duty of due care).
A. Section 102(b)(7) Protects Directors from Liability Arising from Reckless Conduct Because Corporate Recklessness Is a Subset of Gross Negligence

Section 102(b)(7) protects directors from liability arising from recklessness because corporate recklessness is a subset of gross negligence.175 As stated by the Tomczak court and acknowledged by Delaware chancellors and treatises, corporate gross negligence includes reckless indifference.176 Therefore, corporate recklessness is a part of the larger classification of grossly negligent conduct. Because section 102(b)(7) protects directors from liability arising from gross negligence,177 it should also protect them from liability arising from a subset of gross negligence—recklessness.

In Van Gorkom, the Delaware State Supreme Court suggested that recklessness is a subset of gross negligence.178 The Van Gorkom court imposed liability on directors for conduct that the court classified as reckless.179 Although the court described the directors’ conduct as reckless, it ultimately concluded that this reckless conduct did not constitute bad faith.180 Instead, the court held that the directors’ misconduct amounted to gross negligence, which breached the directors’ duty of due care.181

Accordingly, because recklessness is a subset of gross negligence,182 section 102(b)(7) protects directors from personal liability arising from personal liability arising from

176 See Tomczak, 1990 WL 42607, at *12 (quoting Allaun, 147 A. at 261); Rabkin, 547 A.2d at 970; WARD ET AL., supra note 68, § 141.2; Allen et al., supra note 5, at 453.
177 See Malpiede, 780 A.2d at 1092–93 (noting that section 102(b)(7) protects directors from personal liability arising from gross negligence).
178 See Smith v. Van Gorkom, 488 A.2d 858, 871 (Del. 1985) (stating that the lower court erred in finding that grossly negligent directors had not acted “recklessly or improvidently”).
179 See id.
180 See id. at 873.
181 See id.
Personal Liability of Reckless Directors

recklessness. The Delaware State Supreme Court has consistently held that section 102(b)(7) protects directors from liability arising from their own gross negligence. Grossly negligent conduct does not breach the duty of good faith; rather, it breaches the duty of due care. Therefore, once directors establish that the plaintiff’s recklessness claim amounts only to a breach of the duty of due care, the directors may invoke section 102(b)(7) to have the court summarily dismiss the claim.

B. Even if Recklessness Is Not a Subset of Gross Negligence, Section 102(b)(7) Protects Directors from Liability Arising from Recklessness Because Reckless Directors Lack a Bad Faith Motive

Even if recklessness is not a subset of gross negligence, section 102(b)(7) protects directors from liability arising from recklessness. The Delaware State Supreme Court requires proof of an illicit motive or bad faith state of mind to establish that conduct was carried out in bad faith. Because recklessness describes only conduct that lacks intent to cause harm, reckless directors lack the illicit motive or bad faith state of mind intrinsic to bad faith conduct. Without this requisite state of mind, reckless directors do not breach the duty of good faith, but instead breach only the duty of due care. Because recklessness is a breach of the duty of due care, section 102(b)(7) protects directors from personal liability arising from their recklessness.

184. See, e.g., Emerald Partners v. Berlin, 787 A.2d 85, 91 (Del. 2001) (stating that Delaware courts have consistently held that section 102(b)(7) protects directors from liability arising from all duty of due care violations); Malpiede, 780 A.2d at 1092–93 (noting that section 102(b)(7) protects directors from personal liability arising from gross negligence).
185. See, e.g., Van Gorkom, 488 A.2d at 873–74 (concluding that the directors acted with gross negligence but in good faith).
186. See Malpiede, 780 A.2d at 1092–93.
188. See supra Part I.C.3.
189. See Desert Equities, 624 A.2d at 1208.
190. See id.
1. A Breach of the Duty of Good Faith Requires a Bad Faith Motive

The Delaware State Supreme Court requires that defendants have a bad faith motive or illicit intent to breach the duty of good faith. For example, the Desert Equities court, focusing on the defendants’ subjective intent, stated that the plaintiff’s allegations of bad faith relied upon the defendants’ “tortious state of mind.” The court further explained that bad faith “implies the conscious doing of a wrong because of dishonest purpose or moral obliquity; . . . it contemplates a state of mind affirmatively operating with furtive design or ill will.” This discussion of bad faith demonstrates that defendants must have a bad faith motive or illicit state of mind to establish that the defendants engaged in bad faith conduct.

As in Desert Equities, the Delaware State Supreme Court in Zirn evaluated the directors’ motives to determine whether they acted in bad faith. The court held that the directors had not breached their duty of good faith because the directors had no “pecuniary motive” or any “other plausible motive for deceiving the stockholders.” Therefore, the court acknowledged that directors must have an illicit motive for their conduct to amount to a breach of the duty of good faith.

Additionally, Van Gorkom demonstrates that the Delaware State Supreme Court will not find a breach of the duty of good faith without proof that directors acted with a bad faith motive. The Van Gorkom court stated that it need not evaluate the directors’ motives because there was no proof or allegation that the directors acted in bad faith.

---

192. See, e.g., Desert Equities, 624 A.2d at 1208 (stating that a claim of bad faith relies upon a person’s state of mind).
193. Id.
194. Id. (emphasis added).
197. Id. (emphasis added).
198. See id.
200. See id.
Personal Liability of Reckless Directors

court suggested that, had the plaintiff alleged bad faith conduct, the court would be required to evaluate the directors’ motives.201 Therefore, to breach the duty of good faith, directors must have a bad faith motive.202

2. Recklessness Lacks a Bad Faith Motive and Therefore Falls Within the Protections of Section 102(b)(7) as a Breach of the Duty of Due Care

Recklessness, by definition, is conduct without a bad faith motive.203 Reckless directors might be conscious of and indifferent to the fact that their conduct could cause harm, but they do not intend to cause the harm.204 A director with a bad faith motive, on the other hand, intends to cause harm.205 Instead of acting with indifference as to whether a negative result materializes, directors who act in bad faith intend to bring about the resulting harm.206 Therefore, reckless directors do not have a bad faith motive.

Without the requisite bad faith motive, recklessness falls within the protections of section 102(b)(7) as a breach of the duty of due care. Reckless directors do not breach the duty of good faith because they have no bad faith motive.207 Unless directors have a conflict of interest with the corporation or its shareholders, the directors do not breach the duty of loyalty.208 Therefore, the only fiduciary duty that a reckless director breaches is the duty of due care.209 Because recklessness is a breach of the duty of due care, section 102(b)(7) protects directors from personal liability arising from recklessness.210

201. See id.
202. See id.; see also Zirn, 681 A.2d at 1061–62 (stating that the directors acted in good faith because they lacked any motive to deceive or harm the corporation or its shareholders).
204. See id.
206. See Desert Equities, 624 A.2d at 1208.
207. See id.
209. See id. (discussing the triad of fiduciary duties).
3. The Disney, McCall, and Abbott Courts Erroneously Held that Directorial Recklessness Breaches the Duty of Good Faith

The Disney chancellor ignored binding Delaware State Supreme Court precedent by holding that reckless conduct breaches a director’s duty of good faith. The Disney chancellor’s description of how the Disney directors abrogated their oversight responsibility is substantially similar to the Delaware State Supreme Court’s description of the Van Gorkom directors’ conduct. Directors from both cases relied almost entirely on the CEO’s recommendation; they made their decisions with little or no deliberation or factual information; and they failed to satisfy their oversight roles by questioning or even evaluating a material corporate decision. Despite the substantially similar directorial conduct, the Delaware State Supreme Court stated that there was no proof or allegations of a breach of the duty of good faith in Van Gorkom, whereas the Disney chancellor opined that the Disney directors’ conduct amounted to a breach of the duty of good faith. Although the Van Gorkom decision set binding precedent, the Disney chancellor failed to distinguish the two cases or even mention Van Gorkom.

In addition to omitting any reference to Van Gorkom, the Disney chancellor also failed to discuss the Delaware State Supreme Court’s other binding precedent established in Desert Equities and Zirn. The Desert Equities court stated that a plaintiff cannot establish bad faith conduct without proving that the defendant had a dishonest purpose or ill will state of mind. The Disney chancellor did not evaluate the

211. See generally In re Walt Disney Co. Derivative Litig., 825 A.2d 275 (Del. Ch. 2003) (failing to discuss Van Gorkom).
213. Compare Van Gorkom, 488 A.2d at 869, 893, with Disney, 825 A.2d at 281, 287.
214. Compare Van Gorkom, 488 A.2d at 874, with Disney, 825 A.2d at 281, 287.
215. Compare Van Gorkom, 488 A.2d at 869–70, with Disney, 825 A.2d at 281–82.
216. See Van Gorkom, 488 A.2d at 873 (finding no proof of bad faith conduct).
217. Disney, 825 A.2d at 286 (“A fair reading of the new complaint, in my opinion, gives rise to a reason to doubt whether the board's actions were taken honestly and in good faith . . . .”).
218. See generally id. (failing to mention Van Gorkom).
219. See generally id. (failing to mention Desert Equities and Zirn).
Personal Liability of Reckless Directors

directors’ state of mind or distinguish Disney from Desert Equities.221 Moreover, the Zirn court stated that directors who act in bad faith have a pecuniary motive or some other motive for deceiving the shareholders.222 Again, the Disney chancellor failed to evaluate the Disney directors’ motives and failed to distinguish Zirn.223

Like the Disney chancellor, the Sixth Circuit in McCall and the Seventh Circuit in Abbott erroneously found reckless conduct sufficient to prove a breach of the duty of good faith.224 The McCall court cited the Nagy decision to support its basic premise that “[u]nder Delaware law, the duty of good faith may be breached where a director consciously disregards his duties to the corporation, thereby causing its stockholders to suffer.”225

However, the Sixth Circuit’s reliance on this Nagy quote suffers from four flaws. First, the quote is from a footnote in the factual background section of a chancery court opinion.226 If Delaware courts had accepted this legal concept, the concept would appear as a holding or at least as a finding in the legal discussion section of other courts’ opinions. Second, the Nagy chancellor cited no authority to support his conclusion.227 Third, no Delaware court has cited Nagy for the proposition that directors’ conscious disregard of their corporate duties breached the duty of good faith even though Delaware courts have faced this issue many times since Nagy.228 Fourth, the Nagy chancellor made this comment while explaining that the duty of good faith has no legal significance independent of the duty of loyalty.229 Because the Abbott court relied exclusively on McCall to conclude that recklessness breaches the duty of...

221. See generally Disney, 825 A.2d 275 (failing to mention Desert Equities or evaluate the directors’ state of mind).
223. See generally Disney, 825 A.2d 275 (failing to mention Zirn or evaluate the directors’ motives).
224. See In re Abbott Labs. Derivative S’holders Litig., 325 F.3d 795, 811 (7th Cir. 2003); McCall II, 250 F.3d 997, 1001 (6th Cir. 2001).
225. McCall II, 250 F.3d at 1001.
227. See id.
228. For example, Disney involved the same issue as addressed in the Nagy footnote, namely whether recklessness is sufficient to breach the duty of good faith, but the Disney chancellor did not cite to Nagy. See generally Disney, 825 A.2d 275.
229. Nagy, 770 A.2d at 49 n.2 (concluding that the duty of good faith is not independent from the duty of loyalty); see also supra Part II.C (discussing lack of consensus about whether section 102(b)(7) protects directors from personal liability arising from reckless conduct).
good faith under Delaware jurisprudence, the Abbott decision suffers from the same flaws as the McCall decision. Accordingly, the Sixth and Seventh Circuits erroneously interpreted Delaware law to hold that reckless conduct is sufficient to prove a breach of the duty of good faith.

Therefore, those courts that held that recklessness breaches the duty of good faith erred. The courts ignored binding Delaware precedent and misinterpreted Delaware law. Instead, they should have concluded that reckless conduct does not breach the duty of good faith because reckless directors do not intend to cause harm.

In sum, section 102(b)(7) protects directors from personal liability arising from recklessness because recklessness is a breach of the duty of due care instead of a breach of the duty of good faith. The Delaware State Supreme Court has clearly required a bad faith state of mind to establish that directors acted in bad faith. Because reckless directors, by definition, lack a bad faith state of mind, they do not act in bad faith. Therefore, recklessness does not breach the duty of good faith, but instead breaches only the duty of due care. As such, section 102(b)(7) protects directors from personal liability arising from recklessness.

IV. CONCLUSION

Currently, confusion exists over whether section 102(b)(7) protects directors from personal liability arising from their own reckless conduct. When addressing this issue, the Delaware State Supreme Court should resolve the confusion by concluding that section 102(b)(7) protects directors from personal liability arising from recklessness for two reasons. First, recklessness is merely a subset of gross negligence in Delaware corporate jurisprudence. Because section 102(b)(7) unambiguously protects directors from personal liability arising from gross negligence, it concomitantly protects directors from liability arising from recklessness, which is a subset of gross negligence. Second,
Personal Liability of Reckless Directors

the Delaware State Supreme Court requires proof of an illicit motive or bad faith state of mind before concluding that conduct was carried out in bad faith. Because recklessness describes only conduct with no intent to cause harm, reckless directors lack the illicit motive or bad faith state of mind intrinsic to bad faith conduct. Therefore, the Delaware State Supreme Court should remain consistent with Delaware precedent by holding that section 102(b)(7) protects directors from personal liability arising from their own recklessness.

Protecting directors from personal liability arising from their recklessness may seem counterintuitive. However, it is important to remember that section 102(b)(7) protects directors only if shareholders first vote to amend the corporation's certificate of incorporation to include that protection. Further, the Delaware General Assembly specifically acted to provide this protection to directors where shareholders consent. If the General Assembly believes that the statute produces inequities, it can amend or repeal the statute at any time.
LUCKY FOR LIFE: A MORE REALISTIC AND REASONABLE ESTATE TAX VALUATION FOR NONTRANSFERABLE LOTTERY WINNINGS

Kyla C.E. Grogan

Abstract: When a lottery winner dies after receiving only a few annuity payments from the state, the winner’s estate must pay federal estate tax on the balance of the annuity. If the lottery prize is not legally transferable, the winner’s estate cannot sell or pledge the right to future payments in order to generate funds to pay the estate tax. The estate tax value of an annuity is generally based on valuation tables that discount the future payments to present value. However, these valuation tables do not control when they produce an unrealistic and unreasonable result. The United States Courts of Appeals are split as to the proper method of valuing nontransferable lottery winnings. The Ninth and Second Circuits have held that the valuation tables produce an unreasonable and unrealistic valuation result for nontransferable lottery winnings because the tables do not consider lack of marketability. The Fifth Circuit departed from this precedent by holding that the valuation tables were controlling and no marketability discount was appropriate for nontransferable lottery winnings. Although stability and lack of transferability are two factors integral to the lottery winnings’ value for estate tax purposes, none of the circuits have considered both of these factors. This Comment proposes a new approach to valuing nontransferable lottery prizes. First, courts should use the valuation tables to calculate the annuity’s present value. Second, courts should modify the tables’ value with a limited-percentage marketability discount calculated by weighing the stable payment structure of the lottery prize against the legal restrictions on its transfer. This approach gives proper weight to the two most fundamental characteristics of nontransferable lottery prizes.

Lotteries are the most common form of gambling in the United States and the only form of gambling operated as a state government monopoly. The state’s backing gives lottery payments great stability because there is very little risk that the state will default. Despite this stability, a lottery winner’s untimely death may create an estate tax dilemma if the decedent’s estate lies in a state where lottery winnings...
are assignable only to the estate or pursuant to a court order. When a lottery winner dies while receiving winnings in the form of an annuity from a state with such rules, the winner’s estate owes tax on the present value of the future annuity payments. Thus, the estate owes tax on the total remaining balance of the lottery winnings, but can pay the tax using only the annual annuity payments.

The standard method for determining the value of an annuity for estate tax purposes uses valuation tables contained in the U.S. Treasury Regulations and Internal Revenue Service publications. The Treasury Regulations specify certain situations where the tables do not apply. Additionally, courts have held that the tables do not control when they produce an “unrealistic and unreasonable” result.

Recently, three U.S. Courts of Appeals have considered the issue of how to value nontransferable lottery winnings. However, the circuits have used different methods of calculating the value of nontransferable lottery winnings. The Ninth and Second Circuits departed from the

4. A number of states that operate lotteries have such restrictions. See, e.g., DEL. CODE ANN. tit. 29, § 4808 (2003) (allowing assignment only to a decedent’s estate or pursuant to a court order); N.M. STAT. ANN § 6-24-21(C)(5) (Michie 2002) (same); R.I. GEN. LAWS § 42-61-7 (1998) (same); TEX. GOV’T CODE ANN. § 466.406 (Vernon Supp. 2004–2005) (same); see also IDAHO CODE § 67-7437 (Michie 2001) (stating that lottery winnings may be paid to the winner’s estate, but lacking an allowance for assignment pursuant to a court order). Other states have fewer restrictions. See, e.g., CAL. GOV’T CODE § 8880.325(c) (West Supp. 2004) (allowing assignment as collateral); MASS. ANN. LAWS ch. 10, § 28(3) (Law. Co-op. Supp. 2004) (permitting “assignment of prizes for purposes of paying estate and inheritance taxes”).

5. See Shackleford v. United States, 262 F.3d 1028, 1030 (9th Cir. 2001) [hereinafter Shackleford II].

6. See id.


9. Treas. Reg. § 20.7520-3(b)(1)(iii) (as amended in 1995). The Treasury Regulations provide that it is inappropriate to use the tables in circumstances such as when an annuity could be exhausted before the end of the defined payment period. See Treas. Reg. § 20.7520-3(b)(2)(i). In contrast, the rule that the tables are inappropriate when they produce an unrealistic and unreasonable result comes from case law, rather than from the Treasury Regulations. See O’Reilly v. Comm’r, 973 F.2d 1403, 1407 (8th Cir. 1992) (quoting Weller v. Comm’r, 38 T.C. 790, 803 (1962)).

10. O’Reilly, 973 F.2d at 1407 (quoting Weller, 38 T.C. at 803).

11. See Cook v. Comm’r, 349 F.3d 850, 851 (5th Cir. 2003); Estate of Gribauskas v. Comm’r, 342 F.3d 85, 86 (2d Cir. 2003) [hereinafter Gribauskas II]; Shackleford II, 262 F.3d at 1029. Under the “Golsen rule,” the Tax Court must follow the rulings of the circuit court of appeals for the taxpayer’s jurisdiction. See Golsen v. Comm’r, 54 T.C. 742, 757 (1970). However, if the relevant court has not yet ruled on an issue, the Tax Court need not consider rulings from other circuits. See id.

12. Compare Gribauskas II, 342 F.3d at 88–89 (departing from the valuation tables and allowing
Valuation of Nontransferable Lottery Winnings

valuation tables because the tables failed to consider lack of marketability. In contrast, the Fifth Circuit valued nontransferable lottery winnings according to the valuation tables and without applying a discount for lack of marketability.

A valuation approach derived from the balancing process used by the Tax Court and affirmed without published opinion by the U.S. Court of Appeals for the Third Circuit to value privately held stock in *Mandelbaum v. Commissioner* produces a more realistic valuation for nontransferable lottery winnings. When valuing an asset requires a marketability discount, courts have discretion to determine the appropriate percentage for this discount. In *Mandelbaum*, the Tax Court demonstrated how to calculate the appropriate percentage for a marketability discount when valuing privately held stock. The court began with the stipulated value of freely traded stock. The court then weighed factors favoring a high marketability discount against factors favoring a low marketability discount in order to determine the appropriate degree of discount to apply to the stock’s stipulated value.

This Comment argues that modifying the valuation tables’ value with a limited-percentage marketability discount calculated using the *Mandelbaum* balancing approach will produce a more realistic and reasonable valuation for nontransferable lottery winnings than the approach of either side of the current circuit split. Part I of this Comment describes how courts value lottery winnings as private annuities. Part II discusses marketability discounts and the balancing approach the *Mandelbaum* court used to calculate the appropriate degree for a marketability discount for privately held stock. Part III summarizes the cases that led to the current circuit split. Finally, Part IV argues that courts should modify the valuation tables’ value with a limited-percentage marketability discount to produce a more realistic and

---

13. See Gribauskas II, 342 F.3d at 88–89; Shackleford II, 262 F.3d at 1032–33.
16. See, e.g., id. at 2867 (rejecting the experts’ opinions and calculating a new marketability discount).
17. See id. at 2864.
18. See id.
19. See id. at 2864, 2867–69.
reasonable valuation for nontransferable lottery winnings.

I. COURTS USE VALUATION TABLES TO VALUE ANNUITIES WHEN THE TABLES PRODUCE REALISTIC AND REASONABLE RESULTS

All facts affecting an asset’s value are relevant in the estate tax valuation process for the asset. If the Internal Revenue Service (the Service) and a taxpayer disagree about the proper valuation for an asset, courts have considerable discretion in determining the asset’s value. For estate tax valuation purposes, courts treat nontransferable lottery prizes paid over time as private, noncommercial annuities. Valuation tables set forth in the Treasury Regulations provide the standard method for valuing private annuities, except in situations where courts determine that the tables produce an “unrealistic and unreasonable” result.

A. When Parties Disagree About the Estate Tax Valuation of an Asset, Courts Have Discretion to Determine Its Fair Market Value

The estate tax is an excise tax on the transfer of property after the owner dies. The taxable estate generally includes all of the decedent’s property interests at death. The Internal Revenue Code describes how to value the assets comprising the taxable estate. Generally, the fair market value of the asset at the time of the decedent’s death provides the basis for its estate tax valuation.
Valuation of Nontransferable Lottery Winnings

Treasury Regulations define fair market value as “the price at which the property would change hands between a willing buyer and a willing seller, neither being under any compulsion to buy or sell and both having knowledge of relevant facts.” Even a completely nonmarketable asset may have a fair market value. Any factor with the potential to affect the asset’s value is relevant to the fair market value calculation.

When the Service and a taxpayer disagree about the proper valuation for an asset, courts have considerable discretion in determining the asset’s value. Although courts presume that the Service’s assessment of value is correct, a taxpayer may overcome this presumption by producing credible evidence of an alternate valuation. However, if a court disagrees with the expert valuations provided by both the taxpayer and the Service, then the court may perform its own valuation of the asset. The only constraint on the court’s discretion is that the evidence must reasonably support its valuation of the asset. A court thus may take an active role in determining the proper valuation for an asset.

B. Valuation Tables Provide a Standard Method for Valuing Annuities, Unless the Tables Produce an Unrealistic and Unreasonable Result

Courts characterize lottery prizes as private or noncommercial annuities. The most common method of valuing an annuity is the

30. See Bank of Cal. v. Comm’r, 133 F.2d 428, 433 (9th Cir. 1943) (stating “we are required to assume the existence of a willing buyer and a willing seller, regardless of whether they actually existed or not, and to assume that the property could and would change hands, even though such a change could not in fact occur”). Thus, lack of marketability does not mean that an asset has no fair market value. However, lack of marketability may reduce an asset’s fair market value. See Mailloux v. Comm’r, 320 F.2d 60, 62 (5th Cir. 1963); see also infra Part II.A (discussing how a marketability discount reduces an asset’s fair market value).
32. See Anderson v. Comm’r, 250 F.2d 242, 249 (5th Cir. 1957); Estate of Newhouse v. Comm’r, 94 T.C. 193, 217 (1990) (describing a court’s power to accept or reject experts’ opinions).
34. See id.
35. See Anderson, 250 F.2d at 249; Newhouse, 94 T.C. at 217.
36. See Anderson, 250 F.2d at 249 (“It is not necessary that the value arrived at by the trial court be a figure as to which there is specific testimony, if it is within the range of figures that may properly be deduced from the evidence.”).
37. See Cook v. Comm’r, 349 F.3d 850, 855 (5th Cir. 2003); Shackleford II, 262 F.3d 1028, 1031 (9th Cir. 2001).
income capitalization approach. This approach uses a discount rate to calculate the present value of each future payment and adds these present values to determine the annuity’s total present value.

Section 7520 of the Internal Revenue Code explains that the Service uses a series of valuation tables (§ 7520 tables) contained in the Treasury Regulations to determine the value of different assets, including annuities. Different tables apply depending on whether the annuity lasts for a fixed term or for the period of a single life. For annuities that last for a fixed term, the valuation table considers only two factors: the term of years and the applicable interest rate.

Valuation tables do not control when their result is “so unrealistic and unreasonable that either some modification in the prescribed method should be made, or complete departure from the method should be taken.” The Tax Court first articulated this principle in Weller v. Commissioner, and the U.S. Court of Appeals for the Eighth Circuit adopted the principle when disregarding the valuation tables in O’Reilly v. Commissioner. Because the valuation tables significantly undervalued a retained interest in stock shares, the O’Reilly court held that the value produced by the tables was “not anywhere close to” the proper valuation of the retained interest, and remanded the case to the

---

40. See id. at 3-49.
42. See I.R.S. Notice 89-60, 1989-1 C.B. 700.
43. I.R.S. Notice 89-24, 1989-1 C.B. 660; see Cook v. Comm’r, 349 F.3d 850, 854 (5th Cir. 2003).
44. O’Reilly v. Comm’r, 973 F.2d 1403, 1407 (8th Cir. 1992). Although O’Reilly is a gift tax case, the same valuation rules and precedents generally apply to both gift and estate tax cases. See Bogdanski, supra note 40, at 2-159.
45. 38 T.C. 790, 803 (1962).
46. 973 F.2d 1403, 1407 (8th Cir. 1992). Although O’Reilly is a gift tax case, the same valuation rules and precedents generally apply to both gift and estate tax cases. See Bogdanski, supra note 39, at 2-159.
47. O’Reilly, 973 F.2d at 1408.
Valuation of Nontransferable Lottery Winnings

Tax Court to determine its proper value.\textsuperscript{48} Courts have interpreted the Treasury Regulations to allow similar departures from the § 7520 valuation tables.\textsuperscript{49} Annuities are valued according to the § 7520 tables unless another regulation permits departure from the tables.\textsuperscript{50} The general fair market value regulation in the Treasury Regulations permits departure from the valuation tables when the tables do not provide a reasonable and realistic approximation of the fair market value of the annuity.\textsuperscript{51} If the tables are not used to determine the value of an annuity, its value is its fair market value based on the relevant facts and circumstances.\textsuperscript{52} Although courts may depart from the tables, the tables are presumptively correct.\textsuperscript{53} Accordingly, a party seeking to depart from the tables bears the burden of proving that they produce an unrealistic and unreasonable valuation result.\textsuperscript{54}

In sum, courts have significant discretion when determining the proper estate tax valuation for an asset.\textsuperscript{55} For valuation purposes, courts treat nontransferable lottery prizes like private annuities.\textsuperscript{56} The Service instructs taxpayers to use valuation tables contained in the Treasury Regulations to value annuities.\textsuperscript{57} However, when the tables produce an unrealistic and unreasonable result, courts should modify or depart completely from this result when a more realistic and reasonable valuation method is available.\textsuperscript{58} When the valuation tables are not used to value an asset, its valuation depends upon the relevant facts and circumstances of the case.\textsuperscript{59}

\textsuperscript{48} Id. at 1409.

\textsuperscript{49} See Shackleford II, 262 F.3d 1028, 1031 (9th Cir. 2001).

\textsuperscript{50} See id.

\textsuperscript{51} See Shackleford II, 262 F.3d at 1031 (citing Treas. Reg. § 20.2031-1(b) (as amended in 1965)). The Gribauskas court also mentioned this regulation, but departed from the tables solely on the basis of case law. See Gribauskas II, 342 F.3d 85, 87–89 (2d Cir. 2003).

\textsuperscript{52} See Treas. Reg. § 20.7520-3(b)(1)(iii) (as amended in 1995).

\textsuperscript{53} See Shackleford II, 262 F.3d at 1033; O’Reilly, 973 F.2d at 1408–09.

\textsuperscript{54} See Shackleford II, 262 F.3d at 1033.

\textsuperscript{55} See Anderson v. Comm’r, 250 F.2d 242, 249 (5th Cir. 1957); Estate of Newhouse v. Comm’r, 94 T.C. 193, 217 (1990).

\textsuperscript{56} See Cook v. Comm’r, 349 F.3d 850, 855 (5th Cir. 2003); Shackleford II, 262 F.3d at 1031.

\textsuperscript{57} See 26 U.S.C. § 7520 (2000); Treas. Reg. § 20.7520-1(c) (as amended in 2000); see also I.R.S. Notice 89-24, 1989-1 C.B. 660 (providing guidance on how to value annuities according to § 7520).

\textsuperscript{58} O’Reilly, 973 F.2d at 1407 (quoting Weller v. Comm’r, 38 T.C. 790, 803 (1962)).

II. MARKETABILITY DISCOUNTS DECREASE THE VALUE OF ASSETS THAT ARE NOT FREELY TRANSFERABLE

A marketability discount lowers the initial valuation of an asset to account for the fact that an asset with restrictions on its transfer is inherently less valuable in the free market than an asset with no such restrictions. In Mandelbaum, the Tax Court illustrated how to calculate a limited-percentage marketability discount for privately held stock by weighing factors favoring a high marketability discount against factors favoring a low marketability discount. The Tax Court has used a similar approach to calculate limited-percentage marketability discounts for other types of assets.

A. A Marketability Discount Reduces an Asset’s Value to Account for Restrictions on Its Transferability

After determining an asset’s initial value, the next step is to consider whether the valuation warrants a discount and, if so, how large the discount should be. A marketability discount decreases the value of an asset that has restrictions on its transferability or liquidity because a buyer will pay less for a restricted asset than for an identical asset with no such restrictions. The most common method of applying a marketability discount is to follow the standard valuation method appropriate for the asset, and then apply a percentage marketability discount to the preliminary valuation result.

60. See, e.g., Newhouse, 94 T.C. at 249 (lowering the initial valuation of stock for lack of marketability).
61. See Mailloux v. Comm’r, 320 F.2d 60, 62 (5th Cir. 1963); Cooley v. Comm’r, 33 T.C. 223, 225 (1959); see also Youpee v. Babbitt, 67 F.3d 194, 197 (9th Cir. 1995) (describing the right to transfer as an essential property right), aff’d, 519 U.S. 234 (1997).
64. See Newhouse, 94 T.C. at 249.
65. See Blanton, supra note 3, at 447.
66. See Mailloux, 320 F.2d at 62; Cooley, 33 T.C. at 225; see also Campfield et al., supra note 41, at 207 (explaining that “[t]he willing buyer–willing seller approach can produce unrealistically high values if the market is thin, buyers are few, or the asset has unusual infirmities”).
67. See Bogdanski, supra note 39, at 3–74 (describing the marketability discount process for closely held stock); see also Baradin, 72 T.C.M. (CCH) at 492–94 (valuing a partnership interest by first applying a capitalization rate to reduce lease income to present value, then applying minority
Valuation of Nontransferable Lottery Winnings

Courts have discretion to determine the appropriate percentage for a marketability discount. Although parties may rely on experts to calculate their proposed marketability discounts, courts are not bound to give these expert opinions any particular weight. If a court disagrees with the parties’ proposed discounts, it may weigh various factors to determine the appropriate marketability discount for an asset. The Tax Court illustrated this balancing approach in Mandelbaum, a memorandum decision.

B. The Mandelbaum Court Used a Fact-Specific Balancing Approach to Calculate the Appropriate Degree for a Marketability Discount

In Mandelbaum, the Tax Court utilized a balancing approach in order to determine the appropriate marketability discount when valuing privately held stock for gift tax purposes. Members of the Mandelbaum family were the sole shareholders of “Big M,” a privately held corporation. The Tax Court determined that a marketability discount was necessary in valuing the shares because privately held shares are not readily marketable. The marketability of Big M shares was further restricted because shareholders’ agreements prohibited transfers outside

---


70. See, e.g., Desmond, 77 T.C.M. (CCH) at 1533–34 (weighing factors to determine the appropriate discount); Mandelbaum, 69 T.C.M. (CCH) at 2867–69 (weighing factors to determine the appropriate discount). See infra note 82 for the Mandelbaum factors. The Desmond factors were nearly the same, although they included potential environmental liabilities rather than the cost of a public offering and modified the factor comparing the values of publicly and privately held stock to the availability of a public market. See Desmond, 77 T.C.M. (CCH) at 1532–34. The average marketability discount is between fifteen and forty percent. RICHARD B. STEPHENS ET AL., FEDERAL ESTATE AND GIFT TAXATION INCLUDING THE GENERATION-SKIPPING TRANSFER TAX 4-40 (8th ed. 2002).

71. See Mandelbaum, 69 T.C.M. (CCH) at 2852. Memorandum decisions are not binding authority. See Nico v. Comm’r, 67 T.C. 647, 654 (1977), rev’d in part on other grounds, 565 F.2d 1234 (2d Cir. 1977). However, courts often cite them as persuasive authority. See BOGDANSKI, supra note 39, at 1-30. The Service has specifically endorsed the Mandelbaum approach for calculating a marketability discount for closely held stock. See U.S. INTERNAL REVENUE SERV., IRS VALUATION TRAINING FOR APPEALS OFFICERS COURSEBOOK 9-6 (1998).

72. See Mandelbaum, 69 T.C.M. (CCH) at 2867–69.

73. See id. at 2854.

74. See id. at 2864.
the family unless the other family members had a right of first refusal. 75 Although both the Mandelbaum family and the Service proposed marketability discounts, the Tax Court found neither party’s expert convincing. 76 The Tax Court thus exercised its discretion to determine the appropriate marketability discount for the shares of Big M stock, 77 and applied this discount to the stock’s stipulated value. 78

The Tax Court took a two-step approach to determine the fair market value of the privately held shares. 79 The parties first stipulated to the value the stock shares would have if they were freely traded. 80 This figure was the base upon which the court applied the marketability discount to account for the fact that the stock was not freely tradable. 81

The court weighed various factors to determine the appropriate marketability discount as part of the second Mandelbaum step. 82 For each factor, the court determined whether the evaluation was neutral or favored an average, above-average, or below-average discount. 83 Big M’s financial statements, dividend policy, economic outlook, and management were uniformly strong, which favored a low marketability discount. 84 In contrast, the restrictions on transferability of the shares and the high costs of a public offering favored an average to above-average

---

75. See id. at 2856, 2866 (describing the shareholders’ agreements and explaining that the agreements affect value, although this effect is “not necessarily substantial”).

76. See id. at 2866.

77. See id. at 2867–69.

78. See id. at 2869.

79. See id. at 2864; see also Estate of Kaufman v. Comm’r, 77 T.C.M. (CCH) 1779, 1784 (1999) (describing the two-step process to determine the fair market value of privately owned shares), rev’d on other grounds sub nom. Morrissey v. Comm’r, 243 F.3d 1145 (9th Cir. 2001).

80. See Mandelbaum, 69 T.C.M. (CCH) at 2864.

81. See id.

82. See id. at 2867–69. These factors included:

1. The value of the subject corporation’s privately traded securities vis-a-vis its publicly traded securities (or, if the subject corporation does not have stock that is traded both publicly and privately, the cost of a similar corporation’s public and private stock); (2) an analysis of the subject corporation’s financial statements; (3) the corporation’s dividend-paying capacity, its history of paying dividends, and the amount of its prior dividends; (4) the nature of the corporation, its history, its position in the industry, and its economic outlook; (5) the corporation’s management; (6) the degree of control transferred with the block of stock to be valued; (7) any restriction on the transferability of the corporation’s stock; (8) the period of time for which an investor must hold the subject stock to realize a sufficient profit; (9) the corporation’s redemption policy; and (10) the cost of effectuating a public offering of the stock to be valued.

Id. at 2864.

83. See id. at 2867–69.

84. See id. at 2867–68.
Valuation of Nontransferable Lottery Winnings

marketability discount.85 After weighing these factors, the Tax Court rejected the seventy to seventy-five percent marketability discount proposed by the Mandelbaums,86 and arrived at a limited-percentage marketability discount of thirty percent.87

C. The Tax Court Has Used a Flexible Balancing Approach Similar to the Mandelbaum Approach to Calculate the Appropriate Marketability Discount for Other Types of Assets

The Tax Court has used a balancing approach similar to its approach in Mandelbaum to determine the appropriate marketability discount for assets other than privately held stock.88 This balancing approach weighs factors favoring a higher marketability discount against factors favoring a lower marketability discount.89 However, the particular factors considered in each case may change depending on the nature of the asset being valued.90

The Tax Court took a balancing approach when determining the appropriate marketability discount for a partnership interest in Estate of Barudin v. Commissioner,92 and for a trust having a note as its principal asset in Estate of Luton v. Commissioner.93 The assets in both cases had stable payment structures.94 However, despite their stability, the Tax Court held that transfer restrictions on the assets made marketability
discounts necessary in both cases.\textsuperscript{95}

In order to use the Mandelbaum balancing process to determine the appropriate degree for a marketability discount, the taxpayer must provide sufficiently detailed evidence in the record for the court to evaluate.\textsuperscript{96} If the taxpayer does not do so, the court will decline to use the balancing process.\textsuperscript{97} Thus, the Mandelbaum balancing process does not alter the fundamental presumption that the Service’s valuation is correct.\textsuperscript{98}

III. THE FIFTH CIRCUIT CREATED A CIRCUIT SPLIT OVER HOW TO VALUE NONTRANSFERABLE LOTTERY PRIZES

The question of the proper method of valuing a nontransferable lottery prize for estate tax purposes first arose when the U.S. Court of Appeals for the Ninth Circuit affirmed a departure from the § 7520 tables after holding that the tables produced an unrealistic and unreasonable result by failing to consider the lack of marketability of nontransferable lottery winnings.\textsuperscript{99} The Second Circuit agreed.\textsuperscript{100} However, the Fifth Circuit created a circuit split when it held that an estate must use the § 7520 tables to value nontransferable lottery winnings without a marketability discount.\textsuperscript{101}

A. The Ninth and Second Circuits Departed from the Section 7520 Valuation Tables and Allowed a Marketability Discount for Nontransferable Lottery Prizes

The Ninth and Second Circuits confronted similar fact patterns in

\textsuperscript{95} See Barudin, 72 T.C.M. (CCH) at 494; Luton, 68 T.C.M. (CCH) at 1053–55.
\textsuperscript{96} See, e.g., Estate of Kaufman v. Comm’r, 77 T.C.M. (CCH) 1779, 1789–90 (1999) (refusing to apply the Mandelbaum factors because of the insufficient record), rev’d on other grounds sub nom. Morrissey v. Comm’r, 243 F.3d 1145 (9th Cir. 2001); Estate of Scanlan v. Comm’r, 72 T.C.M. (CCH) 613, 615 (1996) (affirming the Tax Court’s prior refusal to apply the Mandelbaum factors because of the insufficient record); Estate of Cloutier v. Comm’r, 71 T.C.M. (CCH) 2001, 2003 (1996) (refusing to apply the Mandelbaum factors when the expert’s report contained “no meaningful discussion of any of the factors of valuation”).
\textsuperscript{97} See Kaufman, 77 T.C.M. (CCH) at 1789–90.
\textsuperscript{99} See Shackleford II, 262 F.3d 1028, 1032–33 (9th Cir. 2001).
\textsuperscript{100} See Gribauskas II, 342 F.3d 85, 88–89 (2d Cir. 2003).
\textsuperscript{101} Cook v. Comm’r, 349 F.3d 850, 857 (5th Cir. 2003).
Valuation of Nontransferable Lottery Winnings

*Shackleford v. United States*¹⁰² and *Estate of Gribauskas v. Commissioner.*¹⁰³ The lottery winners in both cases received their winnings as twenty-year annuities and died after receiving only two and three payments, respectively.¹⁰⁴ In both jurisdictions, the right to receive these lottery winnings was nontransferable.¹⁰⁵ In light of these facts, both circuits ultimately departed from the valuation tables and applied a marketability discount to the nontransferable lottery winnings.¹⁰⁶

In *Shackleford*, the estate paid the tax computed by the § 7520 tables, but later claimed the tables overvalued the winnings and sued the Service for a refund in federal district court.¹⁰⁷ At trial, the estate’s primary valuation expert disregarded the § 7520 tables’ interest rate of 10.4%; instead, he studied the discount rates for other comparable assets in order to arrive at a thirty-five percent discount rate.¹⁰⁸ This gave the lottery prize a value of only $1,900,000,¹⁰⁹ in contrast to the tables’ value of $4,023,903.¹¹⁰

The district court stated that “[t]he ‘reality’ of a decedent’s economic interest” in the asset was a major component of its value,¹¹¹ and accordingly held that the lack of marketability must be considered when valuing the lottery prize for estate tax purposes.¹¹² The court thus concluded that the estate’s departure from the valuation tables was warranted because the tables produced an unrealistic and unreasonable result by failing to consider lack of marketability.¹¹³ However, the court declined to accept the estate’s valuation because the estate did not make adjustments for various other factors, including the lack of risk of the lottery payments.¹¹⁴

The Ninth Circuit affirmed the district court’s conclusion that the

---

¹⁰² 262 F.3d 1028 (9th Cir. 2001).
¹⁰³ 342 F.3d 85 (2d Cir. 2003).
¹⁰⁴ See id. at 86; *Shackleford II*, 262 F.3d at 1030.
¹⁰⁵ See *Gribauskas II*, 342 F.3d at 86; *Shackleford II*, 262 F.3d at 1030.
¹⁰⁶ See *Gribauskas II*, 342 F.3d at 88–89; *Shackleford II*, 262 F.3d at 1032–33.
¹⁰⁷ See *Shackleford II*, 262 F.3d at 1030.
¹⁰⁹ Id. at 11.
¹¹¹ See id. at *4* (citing Helvering v. Safe Deposit Trust Co., 316 U.S. 56, 58 n.1 (1946)).
¹¹² See id. at *5*.
¹¹³ See id.
¹¹⁴ See id. at *3*.
nontransferability of the lottery winnings justified a marketability discount. The court explained that the right to transfer is an important property right, and that the transfer restrictions on the remaining lottery payments accordingly reduced their fair market value. The court noted that the case law authorizes either a modification of the tables’ value or a complete departure from the tables when they produce an unrealistic and unreasonable result that does not reasonably approximate fair market value. The Ninth Circuit affirmed the district court’s conclusion that the tables produced an unrealistic and unreasonable result because the tables did not consider the important element of marketability, and held that this result justified the district court’s departure from the tables.

In *Gribauskas*, the estate valued the lottery winnings using a higher discount rate than the § 7520 rate, and the Service assessed a tax deficiency against the estate. In response, the estate filed a petition in Tax Court contesting the Service’s valuation. The parties disagreed about how to value the lottery annuity. While the Service used the § 7520 discount rate of 9.4%, the estate calculated a discount rate of fifteen percent. The value produced by the § 7520 tables was more than $900,000 over the estate’s valuation. The Service stipulated that the estate’s figure would be correct if the tables were not used. The Tax Court held that the § 7520 tables provided the proper method for valuing the winnings and that the winnings should not receive a marketability discount. The Tax Court emphasized that the valuation tables serve the important function of increasing consistency and efficiency, and that courts have narrowly construed any exceptions to

115. *See* Shackleford II, 262 F.3d 1028, 1032–33 (9th Cir. 2001).
116. *See* id. at 1032.
117. *See* id. at 1031–32.
118. *See* id. at 1032–33.
119. *See* id. at 1029.
120. *See* Gribauskas II, 342 F.3d 85, 86–87 (2d Cir. 2003).
121. *See* id. at 87.
123. *Id.* In arriving at the fifteen percent rate, the expert considered “risk, inalienability, illiquidity, and lack of marketability.” *Id.*
125. *See* id. at 87.
Valuation of Nontransferable Lottery Winnings

their use.\textsuperscript{127}

On appeal, the Second Circuit reversed the Tax Court’s decision.\textsuperscript{128} The Second Circuit recognized that the pre-Shackleford cases in which courts allowed departures from the tables involved situations where there was an inconsistency between a factual assumption of the tables and the facts of the case.\textsuperscript{129} However, the Second Circuit held that the same reasoning that required departures from the tables in those cases also required departures from the tables in cases involving a substantial error in the tables’ ultimate valuation result.\textsuperscript{130} The court explained that the party seeking to depart from the tables has the considerable burden of proving that the tables’ value is unrealistic and unreasonable; this burden preserves the consistency the tables provide in most cases.\textsuperscript{131} Thus, both the Ninth and Second Circuits departed from the strict use of the valuation tables because the tables produced an unrealistic and unreasonable result by failing to consider the lack of marketability of nontransferable lottery winnings.\textsuperscript{132}

B. The Fifth Circuit Held that Nontransferable Lottery Prizes Must Be Valued Using the Section 7520 Tables Without a Marketability Discount

In contrast to the Second and Ninth Circuits, the Fifth Circuit held in \textit{Cook v. Commissioner}\textsuperscript{133} that a marketability discount was inappropriate for nontransferable lottery winnings.\textsuperscript{134} Cook and her sister-in-law won the lottery as joint participants in a partnership devoted to the purchase of lottery tickets.\textsuperscript{135} Cook died after the partnership had received only one annual payment from a twenty-year annuity,\textsuperscript{136} and the state

\begin{footnotesize}
\begin{enumerate}
  \item See id. at 162.
  \item Gribauskas II, 342 F.3d at 86.
  \item See id. at 88. For an example where the interest rate assumption of valuation tables was inconsistent with the facts of the case, see, e.g., Berzon v. Comm’r, 534 F.2d 528, 532 (2d Cir. 1976) (explaining that the taxpayers could not use the tables because they would produce a clearly erroneous valuation for an asset that would produce no income).
  \item See Gribauskas II, 342 F.3d at 89.
  \item See id.
  \item See id. at 88–89; Shackleford II, 262 F.3d 1028, 1032–33 (9th Cir. 2001).
  \item 349 F.3d 850 (5th Cir. 2003).
  \item See id. at 857.
  \item Id. at 852. The partnership element did not alter the core issue of how the court valued the lottery winnings. See id. at 853–54.
  \item Id. at 851–52.
\end{enumerate}
\end{footnotesize}
prohibited assignment of the right to receive the lottery winnings without a court order. Cook’s valuation expert used a discounted cash flow method that incorporated a marketability discount to value the estate’s interest in the partnership at $1,529,749. The Service assessed a deficiency judgment against the estate after determining that the tables’ valuation of the interest was $3,222,919. The Tax Court held that the § 7520 tables provided the appropriate method for valuing the lottery winnings. The estate appealed to the Fifth Circuit. The Fifth Circuit affirmed, holding that the § 7520 tables provided the appropriate method for valuing nontransferable lottery winnings payable to a decedent’s estate. According to the Cook court, the § 7520 tables provide certainty in valuation that is more important than accuracy in individual cases, except in those cases where the tables’ value is unrealistic and unreasonable. The court criticized the Gribauskas and Shackleford courts’ departures from the tables on the grounds that, before Shackleford, courts had departed from the tables only when those tables made incorrect factual assumptions. The Cook court determined that the tables did not make an incorrect factual assumption about marketability, and refused to apply the exception for unrealistic and unreasonable results on the basis of its conclusion that marketability was not an appropriate factor to consider in determining the value of the lottery winnings.

The Cook court opined that a marketability discount is inappropriate for any asset that involves the right to receive a stream of payments that are not affected by market forces. It stated that marketability should be considered only when capital appreciation is relevant to value or when value is otherwise difficult to calculate in the absence of a market

137. Id. at 851.
138. Id. at 852.
139. Id.
140. See id.
141. Id. at 851–52.
142. Id. at 851.
143. Id. at 854 (citing O’Reilly v. Comm’r, 973 F.2d 1403, 1407 (8th Cir. 1992)).
144. See id. at 856.
145. See id. at 856–57. In contrast, Judge Davis argued in his dissenting opinion that "the better rule, as recognized by the 2nd and 9th Circuits, is to consider any factor that affects the annuity’s fair market value, including its nonmarketability." Id. at 859–60 (Davis, J., dissenting).
146. Id. at 856.
Valuation of Nontransferable Lottery Winnings

exchange. The court reasoned that because the § 7520 tables could reduce the remaining annuity payments of the lottery prize to present value, their value was readily ascertainable and a marketability discount was inappropriate.

IV. COURTS SHOULD MODIFY THE SECTION 7520 TABLES’ VALUE WITH A LIMITED-PERCENTAGE MARKETABILITY DISCOUNT

Neither side of the current circuit split correctly values nontransferable lottery winnings because neither side considers both of the factors that are fundamental to their value: the stability of the payments and their lack of legal transferability. An approach derived from Mandelbaum provides a more realistic and reasonable method of valuing nontransferable lottery winnings. This approach begins with the tables’ valuation and weighs stability against lack of transferability to determine the appropriate degree for a limited-percentage marketability discount to modify the § 7520 tables’ value.

A. Both Sides of the Circuit Split Are Incorrect in Their Approaches to Valuing Nontransferable Lottery Winnings

Neither side of the circuit split correctly values nontransferable lottery winnings because neither side considers both of the key elements affecting their value: stability and lack of transferability. The

147. See id. at 857.
148. See id.
149. Compare id. at 856–57 (prohibiting a marketability discount for lack of transferability), with Gribauskas II, 342 F.3d 85 (2d Cir. 2003) (failing to mention any adjustment for stability), and Shackleford II, 262 F.3d 1028 (9th Cir. 2001) (same).
150. See Cook, 349 F.3d at 856–57 (focusing on stability as the primary reason to prohibit a marketability discount); Gribauskas II, 342 F.3d at 88 (focusing on transferability to justify a marketability discount); Shackleford II, 262 F.3d at 1032–33 (same).
152. See infra Part IV.B.
153. Compare Cook, 349 F.3d at 856–57 (prohibiting a marketability discount for lack of transferability), with Gribauskas II, 342 F.3d 85 (failing to mention any adjustment for stability), and Shackleford II, 262 F.3d 1028 (same).
154. See Cook, 349 F.3d at 856–57 (focusing on stability as the primary reason to prohibit a marketability discount); Gribauskas II, 342 F.3d at 88 (focusing on transferability to justify a marketability discount); Shackleford II, 262 F.3d at 1032–33 (same).
Second and Ninth Circuits departed from the valuation tables when they held that a marketability discount was appropriate for nontransferable lottery winnings due to their lack of legal transferability. However, these circuits did not require the stable payment structure of the lottery winnings to limit the degree of this marketability discount. In contrast, the Fifth Circuit held that a marketability discount was unwarranted for the lottery payments due to their stability and predictability, despite their lack of legal transferability. Thus, neither side of the circuit split correctly valued nontransferable lottery winnings because neither side considered both of the fundamental elements of stability and lack of legal transferability as part of the valuation process.

1. The Second and Ninth Circuits’ Approaches Are Incorrect Because They Do Not Require Use of the Section 7520 Tables and Do Not Limit the Degree of the Marketability Discount

The Second and Ninth Circuits erred in valuing nontransferable lottery winnings because they did not explicitly require the use of the valuation tables and did not require any limitation on the marketability discount for these winnings. First, the Second and Ninth Circuits erred by failing to require the use of the valuation tables as a starting point from which to value nontransferable lottery winnings. These tables create efficiency and consistency in most cases, and courts narrowly construe any exceptions to their use. However, use of the tables alone is inappropriate when they produce an unrealistic and unreasonable result. In such a case, modifying the tables’ value produces a more realistic and reasonable result than abandoning the tables completely because the tables perform the important task of reducing the future

155. See Gribauskas II, 342 F.3d at 88–89; Shackleford II, 262 F.3d at 1029, 1033.
156. See generally Gribauskas II, 342 F.3d 85 (failing to mention any adjustment for stability); Shackleford II, 262 F.3d 1028 (same).
157. See Cook, 349 F.3d at 857.
158. See generally Gribauskas II, 342 F.3d 85 (failing to require use of the valuation tables and failing to mention any adjustment for stability); Shackleford II, 262 F.3d 1028 (same).
159. See generally Gribauskas II, 342 F.3d 85 (failing to require use of the valuation tables); Shackleford II, 262 F.3d 1028 (same).
161. See O’Reilly v. Comm’r, 973 F.2d 1403, 1407 (8th Cir. 1992) (quoting Weller v. Comm’r, 38 T.C. 790, 803 (1962)).
Valuation of Nontransferable Lottery Winnings

stream of annuity payments to present value, a necessary step in any lottery valuation case.

Second, while the Second and Ninth Circuits allowed marketability discounts for nontransferable lottery winnings, these circuits erred in ignoring the element of stability that should limit the degree of such marketability discounts. The district court in Shackleford recognized the importance of stability when it refused to accept the estate’s valuation because the estate did not adjust for factors including the lack of risk of the lottery payments. The Tax Court also demonstrated in Mandelbaum that elements of stability limit the appropriate marketability discount for stock. The Tax Court has extended this principle to other types of assets. Thus, requiring the use of the tables as the base from which to apply a marketability discount and limiting the degree of this marketability discount to account for the stability of the lottery payments produces a more realistic and reasonable result than the Second and Ninth Circuits’ approach.

2. The Fifth Circuit’s Approach Is Incorrect Because It Does Not Allow a Marketability Discount to Account for Lack of Transferability When Valuing Nontransferable Lottery Winnings

The Fifth Circuit erred in refusing to allow a marketability discount

163. See Cook v. Comm’r, 349 F.3d 850, 857 (5th Cir. 2003); Edward P. Wojnaroski, Jr., Private Annuities and Self-Canceling Installment Notes A-93 (Tax Mgmt., Inc., Estates, Gifts, & Trusts Portfolio Series 805-2d, 2002) (explaining that “the federal estate value of lottery payments equals the present value of the right to receive periodic annuity payments”).
164. See Gribauskas II, 342 F.3d at 88–89; Shackleford II, 262 F.3d at 1033.
165. See generally Gribauskas II, 342 F.3d 85 (failing to mention any adjustment for stability); Shackleford II, 262 F.3d 1028 (same).
166. See Shackleford I, No. CIV.S.96-1370LKKPAN, 1999 WL 744121, at *3 (E.D. Cal. Aug. 6, 1999), aff’d, 262 F.3d 1028 (9th Cir. 2001). Although the Ninth Circuit affirmed the district court’s decision, see Shackleford II, 262 F.3d at 1032–33, it did not discuss the importance of limiting the marketability discount for lottery winnings due to their stability. See generally Shackleford II, 262 F.3d 1028.
168. See, e.g., Estate of Barudin v. Comm’r, 72 T.C.M. (CCH) 488, 494 (1996) (weighing stable management and distribution history against lack of a public market and restrictions on transfer of a partnership interest); Estate of Luton v. Comm’r, 68 T.C.M. (CCH) 1044, 1053–55 (1994) (weighing low credit risk against restrictions on transferability for a note that was the primary asset of a trust).
for nontransferable lottery winnings.\textsuperscript{169} When the right to receive lottery winnings is not legally transferable, this limitation justifies a marketability discount.\textsuperscript{170} The right to transfer is an important property right, and the legal absence of this right reduces the fair market value of nontransferable lottery winnings accordingly.\textsuperscript{171} Because the § 7520 tables do not consider the element of marketability, they produce an unrealistic and unreasonable result for nontransferable lottery winnings.\textsuperscript{172}

The Fifth Circuit’s rationale—that a marketability discount is inappropriate because the present value of the lottery prize is readily ascertainable\textsuperscript{173}—fails to consider the purpose of a marketability discount.\textsuperscript{174} The purpose of a marketability discount is not simply to determine value that is difficult to ascertain, but to recognize that an asset with limited transferability is inherently less valuable than an asset with no transferability restrictions.\textsuperscript{175} While nontransferable lottery winnings have an intrinsically stable value,\textsuperscript{176} their lack of legal transferability significantly limits this value.\textsuperscript{177} Thus, courts should not value nontransferable lottery winnings without considering the limitation on their transferability.

\textbf{B. Courts Should Modify the Section 7520 Tables’ Value with a Limited-Percentage Marketability Discount to Produce a More Realistic and Reasonable Valuation}

Courts should modify the § 7520 tables’ value with a limited-percentage marketability discount, calculated by weighing stability against lack of transferability, in order to produce a more realistic and

\begin{itemize}
  \item \textsuperscript{169} See Cook v. Comm’r, 349 F.3d 850, 856–57 (5th Cir. 2003) (denying a marketability discount because of the stability and predictability of payments).
  \item \textsuperscript{170} See Shackleford II, 262 F.3d at 1029.
  \item \textsuperscript{171} See id. at 1032–33.
  \item \textsuperscript{172} See id.
  \item \textsuperscript{173} See Cook, 349 F.3d at 856–57 (denying a marketability discount because of the stability and predictability of payments).
  \item \textsuperscript{174} See Cooley v. Comm’r, 33 T.C. 223, 225 (1959) (explaining that marketability is relevant when valuing an asset to account for the fact that an asset is inherently less valuable when its marketability is restricted).
  \item \textsuperscript{175} See id.
  \item \textsuperscript{176} See Cook, 349 F.3d at 855 (noting the low risk of lottery payments and their immunity from market forces).
  \item \textsuperscript{177} See Shackleford II, 262 F.3d at 1032–33.
\end{itemize}
Valuation of Nontransferable Lottery Winnings

reasonable valuation result for nontransferable lottery winnings. Courts have the discretion to modify an unrealistic and unreasonable result of the valuation tables. Because lack of marketability is the factor that makes the tables’ valuation of nontransferable lottery winnings unrealistic and unreasonable, a marketability discount is the appropriate method of modifying the tables’ value. A balancing approach derived from Mandelbaum weights the stability of lottery winnings against restrictions on their transfer in order to create a limited-percentage marketability discount, which modifies the tables’ value and produces a more realistic and reasonable valuation for nontransferable lottery winnings.

1. Courts Should Modify the Section 7520 Tables’ Value When Valuing Nontransferable Lottery Winnings

When valuing nontransferable lottery winnings, courts should use their discretion to modify the § 7520 tables’ value. When the valuation tables produce an unrealistic and unreasonable result, courts have the discretion to modify this result rather than disregarding the tables entirely. Although the § 7520 tables do not consider marketability, they provide the interest rate component necessary to reduce payments over a term of years to present value. Thus, a court should not disregard the tables entirely when valuing nontransferable lottery winnings because the tables are still capable of performing the important function of reducing the stream of future annuity payments to present value.

Before Shackleford, courts departed from the valuation tables in cases where the tables made incorrect factual assumptions. If a fundamental

---

178. See id. at 1031; O’Reilly v. Comm’r, 973 F.2d 1403, 1407 (8th Cir. 1992) (quoting Weller v. Comm’r, 38 T.C. 790, 803 (1962)).

179. See Cook, 349 F.3d at 855–56 (explaining that the only difference between the tables’ valuation and the experts’ valuation of the lottery prize was that the experts used marketability discounts).

180. See infra Part IV.B.3.

181. See O’Reilly, 973 F.2d at 1407 (quoting Weller, 38 T.C. at 803) (noting the court’s discretion to modify the tables’ value).

182. See Shackleford II, 262 F.3d at 1031; O’Reilly, 973 F.2d at 1407 (quoting Weller, 38 T.C. at 803).


184. See Cook, 349 F.3d at 856 (explaining that prior to Shackleford II, courts had departed from the valuation tables only where those tables made incorrect factual assumptions); Gribauskas II, 342 F.3d 85, 88–89 (2d Cir. 2003) (acknowledging that prior to Shackleford II, courts had departed from
assumption of the tables is incorrect, it is reasonable to disregard the tables’ value because the tables are no longer capable of correctly performing their primary function of reducing a stream of future payments to present value. However, this does not hold true when the tables produce an unrealistic and unreasonable result for a reason unrelated to one of their fundamental assumptions.

When the tables produce an unrealistic and unreasonable result for a reason unrelated to one of their fundamental assumptions, they should retain an important role in the valuation process. The § 7520 tables perform the important function of reducing the future payment stream of an annuity to present value. This step is necessary in any lottery valuation case. The only elements the § 7520 tables consider are interest rate and term of years. Thus, if there are no flaws in the tables’ assumptions regarding either of these elements, the tables correctly perform a necessary valuation step by reducing future lottery payments to present value, and there is no reason to disregard their result entirely.

2. A Marketability Discount Is the Appropriate Way to Modify the Section 7520 Tables’ Value Because Lack of Marketability Makes the Tables’ Unmodified Valuation Unrealistic and Unreasonable

The § 7520 tables may produce an unrealistic and unreasonable result for nontransferable lottery winnings because of their failure to account for marketability. Therefore, a marketability discount is the appropriate method to modify this result. Modifying the tables’ value

the valuation tables only where those tables made incorrect factual assumptions, but concluding that the same reasoning was applicable to cases where the tables produce an unrealistic and unreasonable result).

185 See, e.g., Berzon v. Comm’r, 534 F.2d 528, 532 (2d Cir. 1976) (explaining that the taxpayers could not use the tables to value an asset because the tables would produce a clearly erroneous valuation for an asset that would produce no income).

186 See, e.g., Gribauskas II, 342 F.3d at 88–89 (holding that the tables produced an unrealistic and unreasonable result because they did not consider marketability, not because one of their fundamental assumptions was incorrect).

187 See Cook, 349 F.3d at 857 (describing the tables’ role in reducing the stream of future lottery payments to present value).


189 See Cook, 349 F.3d at 857; Woinaroski, supra note 163, at A-93 (“The federal estate value of lottery payments equals the present value of the right to receive periodic annuity payments.”).


191 See Cook, 349 F.3d at 856.
Valuation of Nontransferable Lottery Winnings

through a marketability discount is particularly appropriate because a discount can be applied on top of an existing valuation result; it does not require a new method of valuation.\(^{192}\) The use of the tables to complete the initial valuation step thus does not preclude modification by a marketability discount. The tables discount a future stream of annuity payments to their present value,\(^{193}\) which is a necessary step in valuing the lottery winnings.\(^{194}\) A marketability discount acknowledges that a nontransferable asset is inherently less valuable than a freely transferable asset.\(^ {195}\) Thus, once the tables complete the first step in the valuation process, courts should apply a marketability discount to modify the tables’ value.

3. **The Mandelbaum Balancing Process Calculates an Appropriate Limited-Percentage Marketability Discount by Weighing the Two Fundamental Attributes of Nontransferable Lottery Winnings**

The Mandelbaum approach of balancing factors favoring a high marketability discount against factors favoring a low marketability discount\(^ {196}\) is particularly appropriate for nontransferable lottery winnings because it requires the stability of the lottery winnings to limit the degree of the marketability discount.\(^ {197}\) The stability of the payments and the statutory restrictions on their transfer are the two fundamental factors considered by courts in determining whether to apply a marketability discount to nontransferable lottery winnings for estate tax valuation purposes.\(^ {198}\) Accordingly, balancing these two factors produces a limited-percentage marketability discount that can modify the § 7520

---

\(^{192}\) See Bogdanski, supra note 39, at 3-74 (describing the marketability discount process for closely held stock); see, e.g., Estate of Barudin v. Comm’r, 72 T.C.M. (CCH) 488, 493–94 (1995) (valuing a partnership interest by first applying a capitalization rate to reduce lease income to present value, then applying minority and marketability discounts to the per-unit liquidation value of the interest).


\(^{194}\) See Cook, 349 F.3d at 857.

\(^{195}\) See Mailloux v. Comm’t, 320 F.2d 60, 62 (5th Cir. 1963); Cooley v. Comm’t, 33 T.C. 223, 225 (1959); Blanton, supra note 3, at 447.


\(^{197}\) See id. at 2867–68 (discussing how Big M’s strong financial statements, dividend policy, economic outlook, and management favored a low marketability discount).

\(^{198}\) See Cook, 349 F.3d at 856 (focusing on stability as the primary reason to prohibit a marketability discount); Gribauskas II, 342 F.3d 85, 88 (2d Cir. 2003) (focusing on transferability to justify a marketability discount); Shackleford II, 262 F.3d 1028, 1032–33 (9th Cir. 2001) (same).
tables’ value and produce a more realistic and reasonable valuation for nontransferable lottery winnings.

The Tax Court weighs stability against transfer restrictions to reach an appropriate marketability discount to value various types of assets. In cases where the court has applied this approach, it has reached a limited-percentage marketability discount. Because Tax Court precedent supports weighing stability against transfer restrictions for various types of assets, this process is appropriate for nontransferable lottery winnings as well.

The Mandelbaum approach for determining an appropriate discount for privately held stock is particularly appropriate for nontransferable lottery winnings because the assets share several common elements. In Mandelbaum, there was no recognized public market for shares, which corresponds to lottery winnings that are not legally transferable. While the restrictive shareholders’ agreements in Mandelbaum prevented shares from being transferred outside the family without the right of first refusal, the transfer restrictions on nontransferable lottery winnings prevent them from being legally transferred to anyone. Finally, the stock in Mandelbaum had elements of stability that limited its marketability discount. Because lottery payments are stable, their stability should similarly lower their marketability discount.

For nontransferable lottery winnings, the result from the § 7520 tables...
Valuation of Nontransferable Lottery Winnings

is equivalent to the result from the first of the two Mandelbaum steps.208 The first Mandelbaum step arrives at a base value from which the court can apply an appropriate marketability discount.209 In Mandelbaum, the parties agreed on this base value by stipulating the value of freely traded stock.210 For nontransferable lottery winnings, the § 7520 tables provide the mechanism for discounting the aggregate remaining annuity payments to present value.211 The tables’ value is thus the starting point equivalent to the first Mandelbaum step because it provides a base to which the court can apply a marketability discount.

The second Mandelbaum step, which involves determining the relevant factors and calculating an appropriate marketability discount,212 requires modification in order to value nontransferable lottery winnings. The Mandelbaum court’s nonexclusive list of elements of value included ten factors.213 However, some of the Mandelbaum factors are not suitable for the valuation of a lottery prize because the Mandelbaum case dealt with the valuation of stock.214 Stock issued by a private corporation and lottery winnings paid by a state government are inherently different.215

A majority of the Mandelbaum factors relate to the two fundamental attributes of nontransferable lottery winnings: stability and lack of legal

208. For an explanation of why the tables should be used initially, see supra Part IV.B.1.
209. See Mandelbaum, 69 T.C.M. (CCH) at 2864.
210. See id.
211. See Cook, 349 F.3d at 857.
212. See Mandelbaum, 69 T.C.M. (CCH) at 2864; see also Estate of Kaufman v. Comm’r, 77 T.C.M. (CCH) 1779, 1784 (1999) (describing the two-step process), rev’d on other grounds sub nom. Morrissey v. Comm’r, 243 F.3d 1145 (9th Cir. 2001).
213. Mandelbaum, 69 T.C.M. (CCH) at 2864; see also supra note 82 (listing Mandelbaum factors).
214. See Mandelbaum, 69 T.C.M. (CCH) at 2864.
215. See id. at 2864 (listing the ten factors the Tax Court considered in valuing Big M’s stock). For example, the factor comparing the difference in value between publicly and privately traded stock does not apply to nontransferable lottery winnings because there is no equivalent to public and private markets for assets that are not legally transferable. A “gray market” for these winnings at a significant discount may exist. Shackleford I, No. CIV.S.96-1370LKKPAN, 1999 WL 744121, at **2–3 (E.D. Cal. Aug. 6, 1999), aff’d, 262 F.3d 1028 (9th Cir. 2001); see Gribauskas II, 342 F.3d 85, 86 (2d Cir. 2003). However, the Tax Court held that this market is irrelevant where the winner cannot legally transfer an enforceable right to payment during his or her lifetime. See Gribauskas I, 116 T.C. 142, 164 (2001), rev’d, 342 F.3d 85 (2d Cir. 2003). Factors involving the degree of control of the corporation transferred with the stock, the period of time for which an investor must hold stock to realize a profit, and the cost of a public stock offering, Mandelbaum, 69 T.C.M. (CCH) at 2864, likewise have no meaningful equivalent in the valuation of a nontransferable lottery prize.
transferability.\textsuperscript{216} Four of the Mandelbaum factors relate to the stability of the asset being valued: the analysis of the corporation’s financial statements; its dividend-paying capacity and history; the nature, history, and economic outlook of the corporation; and its management.\textsuperscript{217} Two Mandelbaum factors relate to impediments to the transfer of the asset being valued: restrictions on transferability\textsuperscript{218} and the corporation’s redemption policy.\textsuperscript{219} Thus, for the valuation of nontransferable lottery winnings, the Mandelbaum factors can be distilled to a two-factor balancing test that focuses on only the two fundamental elements of stability and lack of legal transferability.

This approach furthers the policies of consistency and judicial economy because it increases the burden on the taxpayer, not the burden on the courts.\textsuperscript{220} The estate is responsible for submitting sufficient evidence to support the factors favoring a marketability discount for the court to evaluate.\textsuperscript{221} Otherwise, the Mandelbaum approach is inapplicable,\textsuperscript{222} and the unmodified tables remain a reasonable default solution.

V. CONCLUSION

The current circuit split over the estate tax valuation of nontransferable lottery winnings resulted from each circuit’s focus on only one of the two equally important characteristics of these prizes—stability and lack of marketability. The Mandelbaum approach allows courts to weigh both of these characteristics to achieve an appropriately limited marketability discount. Courts should then apply this discount to the present value of the lottery winnings calculated by the § 7520 tables.

\textsuperscript{216} See Mandelbaum, 69 T.C.M. (CCH) at 2864.
\textsuperscript{217} See id. All of these factors will likely enhance the value of nontransferable lottery winnings because state-funded lottery payments are quite stable. See Cook v. Comm’r, 349 F.3d 850, 855 (5th Cir. 2003).
\textsuperscript{218} See Mandelbaum, 69 T.C.M. (CCH) at 2864.
\textsuperscript{219} Id. This factor is relevant to transferability because the winners of nontransferable lottery winnings paid in the form of an annuity in these cases had no ability to “redeem” the right to receive the annuity payments in exchange for a lump sum payment. See Cook, 349 F.3d at 851; Gribovskas II, 342 F.3d at 86; Shackleford I, 1999 WL 744121, at *1.
\textsuperscript{221} See Kaufman, 77 T.C.M. (CCH) at 1789–90 (1999); Scanlan, 72 T.C.M. (CCH) at 615.
\textsuperscript{222} See Kaufman, 77 T.C.M. (CCH) at 1789–90.
Valuation of Nontransferable Lottery Winnings

This approach allows courts to modify unrealistic and unreasonable results of the § 7520 tables by applying a limited-percentage marketability discount that gives appropriate weight to the two fundamental characteristics of nontransferable lottery winnings.
PUTTING FLESH ON THE BONES OF UNITED STATES V. WINANS: PRIVATE PARTY LIABILITY UNDER TREATIES THAT RESERVE ACTUAL FISH FOR THE TRIBAL TAKING

Lindsay Halm

Abstract: One hundred years ago, in United States v. Winans, the United States Supreme Court announced that private parties are subject to the rights reserved by Indians under treaty. Accordingly, tribes enforce their treaty fishing rights in federal court to halt private and government actions that threaten to impair their reserved right to take a fair portion of fish from usual and accustomed fishing stations. In addition to injunctive relief, federal courts may award monetary relief to tribes where Congress limits the treaty fishing right. In general, monetary relief is a remedy against any defendant actor who impairs non-fishing treaty-reserved rights. Furthermore, courts have long awarded damages to commercial fishers for interference with their vocational rights. Courts in the Ninth Circuit, however, have denied monetary relief to tribes when private projects destroy the treaty right to take fish. This Comment argues that courts should award damages to tribes when private projects proximately cause harm to a tribe’s right to take actual fish.

The right to resort to the fishing places . . . was a part of larger rights possessed by the Indians, upon the exercise of which there was not a shadow of impediment, and which were not much less necessary to the existence of the Indians than the atmosphere they breathed.1

The right to harvest fish is central to many tribes’ existence, culture, and welfare.2 It has existed from time immemorial and continues into perpetuity.3 Notwithstanding the reservation of fishing rights under treaties with the United States, the federal government has listed numerous fish species as threatened or endangered4 under the

3. See Fishing Vessel, 443 U.S. at 667 (interpreting the Yakima Tribe’s understanding that they “would forever be able to continue” fishing practices under treaty).
4. See, e.g., Endangered and Threatened Species; Threatened Status for Three Chinook Salmon Evolutionarily Significant Units (ESUs) in Washington and Oregon, and Endangered Status for One Chinook Salmon ESU in Washington, 64 Fed. Reg. 14308-01 (1999) (listing the Puget Sound,
Endangered Species Act following decades of habitat destruction, dam building, and over-fishing. For example, the Columbia River salmon runs, once the largest in the world, have diminished by seventy-five to eighty-five percent due to the dozens of dams that currently impede fish passage. Despite successful suits by Northwest Tribes to secure a fair portion—up to fifty percent—of the available harvest, various fish species continue to decline rapidly, which might suggest that a fair portion of what is available today may soon be worthless. Said another way, the right to half of zero . . . is still zero.

Federal courts, however, must reconcile the dire warning that tribal fishing rights are doomed to nothingness with precedent that contemplates actual fish for the tribal taking. One hundred years ago, the United States Supreme Court in United States v. Winans declared that private landowners are subject to the treaty fishing right and that accommodation is required to ensure continuing exercise of that right. Failure to uphold the treaty right, the Court stated, results in “an impotent outcome to negotiations and a convention, which seemed to promise more and give the word of the Nation for more.”

Because the U.S. Supreme Court has likened Indian treaties to

---


8. See Fishing Vessel, 443 U.S. at 685–86.

9. See United States v. Washington, 759 F.2d 1353, 1356–57 (9th Cir. 1985) (vacating a portion of the district court opinion on the “environmental issue” that declared that the right to take fish necessarily includes the right to have those fish protected from man-made despoliation).

10. “Actual fish” refers to tangible fish that, reduced to possession, can be used to sustain the needs of a given tribe. See infra Part IV.A.

11. See Fishing Vessel, 443 U.S. at 678; Northwest Res., 35 F.3d at 1377.


13. Id. at 380–81.

14. Id. at 380.
Private Party Liability Under Indian Treaties

contracts between sovereigns, only Congress has the power to limit rights contained therein. Thus, government and private projects that threaten a tribe’s exercise of fishing rights reserved under treaty cannot proceed without express authorization from Congress. In recent decades, federal courts have enjoined both governmental and private projects to protect tribes’ right to harvest fish. Treaty tribes have also secured monetary relief where Congress has limited, or abrogated, fishing rights reserved under treaty. Similarly, where private or government actors interfere with other treaty-reserved rights, such as land or mineral rights, courts grant monetary relief to tribes. In contrast, courts within the Ninth Circuit have rejected monetary relief for the impairment of treaty fishing rights by private projects.

This Comment argues that courts should award monetary relief to tribes when private parties impair tribal treaty rights to take actual fish. As Winans and its progeny indicate, Indian treaties operate as a preexisting legal condition on the landscape, which binds the federal

15. Fishing Vessel, 443 U.S. at 675.
18. Winans, 198 U.S. at 380 (enjoining a private project); Muckleshoot v. Hall, 698 F. Supp. 1504, 1517 (W.D. Wash. 1988) (same). Because the federal government is the only party that maintains a special trust relationship with tribes, the term “private” in this Comment refers to any non-federal projects, including those of municipal corporations.
19. See infra Part II.A (discussing case law that awards relief to tribes in the absence of congressional authorization).
24. See infra Part IV.
government, states, and private citizens. Tribes currently may assert their fishing right before a private party proceeds with project construction; tribes should also be able to seek monetary relief in federal courts after private parties implement harmful projects. When a private project violates a treaty, a court should make the tribe whole by awarding monetary relief as calculated by the proximately caused loss of a fair portion of the fish harvest.

Part I of this Comment reviews the scope and enforceability of treaty fishing rights. Part II discusses the relief granted when governmental and private parties interfere with fishing and other rights reserved under treaty. Part III examines claims for monetary relief against private parties who interfere with fishing rights within the Ninth Circuit. Lastly, Part IV argues that United States v. Winans and its progeny reserve actual fish for the tribal taking; thus, where private parties harm this right, courts should award monetary relief to make a tribe whole.

I. BY TREATY, TRIBES RESERVED THE RIGHT TO ACCESS USUAL AND ACCUSTOMED PLACES AND TAKE FISH

Occupying a unique niche in U.S. Supreme Court jurisprudence, Indian treaties trigger rules of construction that unmistakably favor tribal rights. To date, the Court has interpreted the fishing clause that appears in several Pacific Northwest treaties to include both a tribal right to access “usual and accustomed places” and a “right of taking fish” in common with other nontreaty citizens. Neither governmental nor


27. See infra Part IV.B.

28. See infra Part IV.B; see also Memorandum from the Associate Solicitor, Indian Affairs, to Solicitor 1 (May 25, 1982) (outlining elements of fish damage claims against private parties).


30. Fishing Vessel, 443 U.S. at 674 (discussing the right of taking fish); Seufert Bros. v. United States, 249 U.S. 194, 199 (1919) (discussing the right of access); Winans, 198 U.S. at 384 (discussing both the right of access and the right of taking fish).

Private Party Liability Under Indian Treaties

private parties have authority to limit treaty fishing rights; only Congress can abrogate or modify the terms.

A. Federal Courts Liberally Construe Treaties that Reserve Tribal Fishing Rights

When tribes granted land to the United States by treaty, they reserved the traditional right to hunt and fish. For example, each of the nine Stevens Treaties of the Washington territory read, with scant variation: “[t]he right of taking fish at all usual and accustomed places, in common with all citizens of the Territory . . . [is secured to said Indians].”

Tribes in other parts of the country similarly reserved in their respective treaties the right to hunt and fish. Indeed, these historic rights persist, even if not explicitly stated under treaty.

To determine the scope of rights contained in Indian treaties, the U.S. Supreme Court employs unique canons of construction to account for the circumstances of historical treaty negotiations. The canons instruct courts to construe terms liberally in favor of establishing Indian rights, resolve ambiguities in favor of protecting tribal interests, and interpret provisions as Indians would have naturally understood them at the time of the treaty’s signing. Though the full scope of treaty fishing rights

32. See Winans, 198 U.S. at 384; Muckleshoot v. Hall, 698 F. Supp. 1504, 1514 (W.D. Wash. 1988) (“The federal, City and private defendants here do not have the ability to qualify or limit the Tribes’ geographical treaty fishing right.”).


34. Winans, 198 U.S. at 377–78 (quoting from the Yakima Treaty); id. at 381 (“[T]he treaty was not a grant of rights to the Indians, but a grant of rights from them—a reservation of those not granted.”).

35. Id. at 378.

36. See, e.g., Mille Lacs Band of Chippewa Indians v. Minnesota, 853 F. Supp. 1118, 1122 (D. Minn. 1994) (interpreting a treaty guaranteeing “the privilege of hunting, fishing and gathering the wild rice upon the lands, the rivers and the lakes included in the territory ceded”), aff’d, 124 F.3d 904 (8th Cir. 1997), and aff’d, 526 U.S. 172 (1999).


41. See Fishing Vessel, 443 U.S. at 676 (citing Jones v. Meehan, 175 U.S. 1, 11 (1899)).
remains untested, the U.S. Supreme Court has placed a “broad gloss” on tribal fishing rights.

B. Federal Courts Recognize Two Treaty-Reserved Rights: The Right to Access Usual and Accustomed Places and the Right to Take Fish

Indian treaty fishing rights include a “geographic right,” or the right to access “usual and accustomed grounds and stations,” both on and off reservation land. In the 1905 landmark Winans case, the U.S. Supreme Court first construed the scope of the access right against a private party. The Winans, upstream landowners, held claim to shore land along the Columbia River under patents from the United States. The Court interpreted the Yakima Treaty as running against the United States and its grantees; the treaty therefore survives the subsequent private acquisition of federal lands. The Court reasoned that the right to access fishing stations established in the land an easement—a “servitude upon every piece of land”—enabling the Tribe’s continual exercise of its

42. See United States v. Washington, 759 F.2d 1353, 1357 (9th Cir. 1985) (vacating part of the district court opinion because of insufficient factual support for declaratory judgment that the right to take fish necessarily includes the right to have those fish protected from man-made despoliation). The scope of the treaty fishing right remains uncertain in the Ninth Circuit following United States v. Washington; several commentators have argued for an expansive reading of treaty rights on various theories. See, e.g., Michael C. Blumm & Brett M. Swift, The Indian Treaty Piscary Profit and Habitat Protection in the Pacific Northwest: A Property Rights Approach, 69 U. COLO. L. REV. 407, 412 (1998) (arguing that courts should consider that the treaty fishing right includes a “habitat right”); O. Yale Lewis III, Treaty Fishing Rights: A Habitat Right as Part of the Trinity of Rights Implied by the Fishing Clause of the Stevens Treaties, 27 AM. INDIAN L. REV. 281, 304–11 (2002–03) (arguing that courts should consider that the treaty fishing right includes a “habitat right”); Brian J. Perron, Note, When Tribal Treaty Fishing Rights Become a Mere Opportunity to Dip One’s Net into the Water and Pull It out Empty: The Case for Money Damages when Treaty-Reserved Fish Habitat Is Degraded, 25 WM. & MARY ENVTL. L. & POL’Y REV. 783, 799–803 (2001) (arguing that courts should provide a remedy for habitat destruction); Allen H. Sanders, Damaging Indian Treaty Fisheries: A Violation of Tribal Property Rights?, 17 PUB. LAND & RESOURCES L. REV. 153, 154 (1996) (arguing that the treaty fishing right is a compensable property interest); Mary Christina Wood, The Tribal Property Right to Wildlife Capital (Part II): Asserting a Sovereign Servitude to Protect Habitat of Imperiled Species, 25 VT. L. REV. 355, 359 (2001) (arguing that tribes maintain a property right as a sovereign entity to protect habitat).

43. See Fishing Vessel, 443 U.S. at 679.

44. See Seufert Bros. v. United States, 249 U.S. 194, 199 (1919); United States v. Winans, 198 U.S. 371, 381–82 (1905). Hereinafter, the term “usual and accustomed places,” as it appears in the Stevens Treaties, is referred to as either “fishing stations” or “historic grounds.”

45. Winans, 198 U.S. at 371.

46. Id. at 379.

47. Id. at 381–82.
Private Party Liability Under Indian Treaties

right.\textsuperscript{48} In addition, the U.S. Supreme Court has interpreted the treaty fishing clause to include a separate-but-related “right to take fish.”\textsuperscript{49} In \textit{Winans}, farmers had obtained a license from the State of Washington to operate fish wheels, contraptions that effectively gave them exclusive possession of the fishery.\textsuperscript{50} The Court rejected the argument that Indian treaty rights could be excluded by a state-licensed device and remanded the case to determine an appropriate “adjustment and accommodation” of the harvest between the Winans and the Yakimas.\textsuperscript{51}

In \textit{Washington v. Washington State Commercial Passenger Fishing Vessel Ass’n},\textsuperscript{52} the Court held that the right to take fish in common with all citizens of the Territory reserved to tribes up to fifty percent of the total harvest.\textsuperscript{53} In a six-to-three decision, the \textit{Fishing Vessel} majority construed the fishing right in no uncertain terms: “In our view, the purpose and language of the treaties are unambiguous; they secure the Indians’ right to take a share of each run of fish that passes through tribal fishing areas.”\textsuperscript{54} Citing undisputed evidence of historically abundant and reliable fish runs, the Court concluded that both parties to the treaty had no doubt that the signatory Indians would continue to take as many fish as they needed.\textsuperscript{55} Indeed, the Tribes assented to cede and peacefully grant millions of acres of land precisely because Washington Territory Governor Stevens recognized that the Tribes reserved, into perpetuity, life-sustaining fish.\textsuperscript{56} Governor Stevens avowed to the signatory Tribes, “[t]his paper secures your fish.”\textsuperscript{57}

The Court held that the treaties secured a tribal catch as necessary to

\begin{itemize}
\item \textsuperscript{48} \textit{Id.} at 381, 384.
\item \textsuperscript{49} See \textit{Fishing Vessel}, 443 U.S. 658, 674 (1979); \textit{Winans}, 198 U.S. at 382, 384.
\item \textsuperscript{50} \textit{Winans}, 198 U.S. at 382.
\item \textsuperscript{51} \textit{Id.} at 382, 384.
\item \textsuperscript{52} 443 U.S. 658 (1979).
\item \textsuperscript{53} \textit{Id.} at 686–88; \textit{see also} United States v. Washington, 384 F. Supp. 312, 343–48 (W.D. Wash. 1974) [hereinafter \textit{Boldt}] (holding that treaty tribes are entitled to a fair portion of the harvestable fish), aff’d, 520 F.2d 676 (9th Cir. 1975). The case is uniformly referred to as the “\textit{Boldt}” decision in reference to the name of the federal district court judge who authored the opinion. See Ed Goodman, \textit{Protecting Habitat for Off-Reservation Tribal Hunting and Fishing Rights: Tribal Comanagement as a Reserved Right}, 30 ENVTL. L. 279, 289 n. 42 (2000).
\item \textsuperscript{54} \textit{Fishing Vessel}, 443 U.S. at 679.
\item \textsuperscript{55} \textit{Id.} at 675–76.
\item \textsuperscript{56} \textit{See id.} at 667, 675–77.
\item \textsuperscript{57} \textit{See id.} at 667 n.11.
\end{itemize}
provide for a livelihood or “moderate living.” In doing so, the Court summarily rejected the State and commercial fishers’ assertion that the “in common with” clause of the treaty promised only equal opportunity to fish. An equal opportunity fishing right, which resulted in the Tribes’ paltry two percent of the catch prior to Fishing Vessel, was not only categorically inadequate, but a derision of treaty negotiations that reserved to Tribes a meaningful compensation for the millions of acres they peacefully ceded. Moreover, the Court reasoned that it was inconceivable that either party would have agreed to crowd the Tribes out of their fishing rights to accommodate future settlers. Hence, the Court interpreted the intent of the signatory parties as reserving to the Tribes an enforceable right to “take” an actual, fair portion of fish, not “merely the chance, shared with millions of other citizens, occasionally to dip their nets.” The commercial harvest allocation was reasoned to subsume the Tribes’ existing ceremonial and subsistence needs, though the Court recognized the possibility that future adjustment would be required if the fifty-percent divide did not accommodate such purposes.

C. Federal Courts Enforce Treaty Fishing Rights Against Governmental and Private Parties

Neither private parties nor government actors may undertake actions that reduce or eliminate a tribe’s treaty fishing right. Indeed, a court’s limitation of a treaty right is reversible error. Congress alone has the

---

58. Id. at 670–71, 686–88.
59. Id. at 676–78.
60. Id. at 676–77 n.22.
61. Id.
62. Id. at 678–79.
63. Id. at 688 (“We need not now decide whether priority for such ceremonial and subsistence uses would be required in a period of short supply in order to carry out the purposes of the treaty.”).
64. Puyallup Tribe v. Dep’t of Game, 391 U.S. 392, 398 (1968) (stating the treaty fishing right “may, of course, not be qualified by the state”); Muckleshoot v. Hall, 698 F. Supp 1504, 1514 (W.D. Wash. 1988) (“The federal, City and private defendants here do not have the ability to qualify or limit the Tribes’ geographical fishing right (or to allow this to occur through permits) by eliminating a portion of an Indian fishing ground for a purpose other than conservation.”).
65. See United States v. Washington, 157 F.3d 630, 650 (9th Cir. 1998) (reversing trial court decision where the lower court had improperly limited the treaty right); Cree v. Waterbury, 78 F.3d 1400, 1405 (9th Cir. 1996) (remanding for a full investigation into the historical context at the time of treaty signing where the district court had summarily assumed a treaty highway right as analogous to a previously litigated fishing right).
power to abrogate or limit treaty fishing rights. Moreover, Congress’s abrogation of a treaty fishing right must be express and specific. The only exception to the rule is a narrow one in which states may issue neutral regulations pursuant to a “conservation necessity.”

Even if Congress approves funding for a government project, such approval does not amount to express abrogation of the treaty right. For example, in Confederated Tribes of the Umatilla Indian Reservation v. Alexander, the U.S. District Court for the District of Oregon refused to infer congressional abrogation from general project authorization, and granted declaratory relief to the Tribes. The court found that, if constructed, the federal agency’s dam would prevent wild fish from swimming upstream to spawn and would destroy access to some of the Tribes’ fishing stations by flooding them with up to two hundred feet of water. Because Congress had authorized the dam without apparent knowledge of such impacts, the federal agency’s action constituted an unauthorized, actual taking of fishing rights. Notably, the dam was never constructed.

Just as federal projects cannot qualify a tribe’s treaty fishing right, the Fishing Vessel Court echoed, in accord with Puyallup Tribe v. Department of Game, that governments cannot regulate away treaty rights. Because the salmon harvest proved at once lucrative and diminishing, state agencies leading up to Fishing Vessel attempted to exclude tribal fishers through regulations that favored non-Indian

---

66. See Confederated Tribes of Umatilla Indian Reservation v. Alexander, 440 F. Supp. 553, 555 (D. Or. 1977) (citing Menominee Tribe of Indians v. United States, 391 U.S. 404, 413 (1968)). Abrogation occurs where Congress expressly legislates to eradicate or otherwise alter the terms of a treaty. See, e.g., United States v. Dion, 476 U.S. 734, 738–39 (1986) (requiring express abrogation of possessory land title); Boldt, 520 F.2d 676, 693 (9th Cir. 1975) (noting that “[o]nce a tribe is determined to be a party to a treaty, its rights under that treaty may be lost only by unequivocal action of Congress”).


68. See Puyallup, 391 U.S. at 398.

69. See Umatilla, 440 F. Supp. at 555.


71. Id. at 555 (citing Menominee, 391 U.S. at 413).

72. Id. at 555–56.

73. Id.

74. See Blumn & Swift, supra note 42, at 465.

75. 391 U.S. 392 (1968).

commercial operations. Tribal members who exercised their treaty rights were subject to harassment, violence, and often arrest in an all-out “fish war.” Though the Tribes prevailed in federal court, the Washington State Supreme Court, employing its own interpretation of the area’s historic Indian treaties, upheld state regulations in defiance of federal orders. On appeal to the U.S. Supreme Court, the majority and dissenting justices agreed that it is the federal courts’ duty and province to construe an Indian treaty.

In sustaining federal jurisdiction in *Fishing Vessel*, the Court enforced the treaty fishing right against the state and against private party fishers. The enforceability of treaties against private parties, however, long predates *Fishing Vessel*. Summarizing the relevant precedent, the *Fishing Vessel* Court recognized that it stood on the shoulders of *United States v. Winans*: “The purport of our cases is clear. Nontreaty fishermen may not rely on property law concepts, devices such as the fish wheel, license fees, or general regulations to deprive the Indians of a fair share of the relevant runs of anadromous fish in the case area.” In *Winans*, the private farmers could not employ an absolute land title or a state fish wheel license to trump the Yakima Tribe’s treaty rights. Instead, the Winans’s exercise of rights was “subject to the treaty” just

---

77. See id. at 669–74; *Puyallup*, 391 U.S. at 398 (rejecting a state’s attempt to qualify or limit the treaty fishing right, except by conservation necessity); Tulee v. Washington, 315 U.S. 681, 684 (1942) (rejecting a state’s attempt to charge a license fee as an unlawful limitation of the treaty fishing right).
78. See *Fishing Vessel*, 443 U.S. at 674; Perron, supra note 42, at 792 n.69 (citing Alex Tizon, 25 Years After the Boldt Decision—the Fish Tale that Changed History, SEATTLE TIMES, Feb. 7, 1999, at A1).
79. See *United States v. Washington*, 459 F. Supp. 1020, 1129–30 (W.D. Wash. 1978), aff’d, 573 F.2d 1123 (9th Cir. 1978) (upholding federal district court orders in the face of state regulations); *Boldt*, 384 F. Supp. 312, 343 (W.D. Wash. 1974) (holding that treaty Tribes are entitled to a fair portion of the harvestable fish), aff’d, 520 F.2d 676 (9th Cir. 1975).
81. See *Fishing Vessel*, 443 U.S. at 693–96; id. at 707 (Powell, J., dissenting) (“To be sure, if it were necessary to construe the treaties to produce these results, it would be our duty so to construe them.”).
82. See id. at 676–77.
83. See United States v. Winans, 198 U.S. 371, 381, 384 (1905).
85. Id. at 684.
86. See *Winans*, 198 U.S. at 381, 384.
Private Party Liability Under Indian Treaties

as it was “to the other laws of the land.” 87 Similarly, in Muckleshoot v. Hall, 88 the U.S. District Court for the Western District of Washington reasoned that permitting the private project at issue would, in effect, determine the “time and manner of [tribal] fishing” and the “size of the take”—a power reserved to Congress and, more narrowly, to states regulating under a conservation necessity. 89 Thus, the court summarized, “[t]he federal, City and private defendants here do not have the ability to qualify or limit the Tribes’ geographical treaty fishing right (or to allow this to occur through permits) by eliminating a portion of an Indian fishing ground.” 90

In sum, courts liberally construe treaty fishing rights by interpreting treaties in favor of tribal interests. 91 Courts reject arguments that interpret fishing rights as a mere opportunity to pursue a catch; rather, treaties reserve to tribes the right to access all historic grounds in order to take a fair portion of the harvest from historic stations. 92 The U.S. Supreme Court has determined that the fifty-percent apportionment to tribes and the accompanying easement are consistent with treaty negotiations that guarantee continuing cultural and economic vitality. 93 As a consequence, courts enforce the treaty fishing right against states, the federal government, and private parties. 94

II. TRIBES SECURE INJUNCTIVE AND MONETARY RELIEF TO PROTECT TREATY RIGHTS IN FEDERAL COURT

Tribes secure injunctive and declaratory relief in federal court to defend their treaty fishing rights against governmental and private projects that, if developed, would interfere with the exercise of fishing rights. 95 Additionally, because the treaty fishing right includes a property interest, courts award tribes just compensation where Congress

87. Id. at 382.
89. Id. at 1512 (citing Puyallup, 391 U.S. at 398).
90. Id. at 1514.
93. Id.
94. See Puyallup, 391 U.S. at 398; Muckleshoot, 698 F. Supp. at 1514.
abrogates a fishing right. Other rights secured under treaty are likewise entitled to monetary relief, whether interference with that right results from federal government taking, state interference, or private party action. Moreover, courts have long afforded monetary relief to non-Indian fishers at common law.

A. Federal Courts Enjoin Any Project that Threatens a Tribe’s Treaty Fishing Right

Like treaties with foreign nations, Indian treaties operate as the supreme law of the land. Consequently, absent express congressional enactment dictating otherwise, courts enjoin private and government projects that, if constructed, would impair treaty fishing rights. Indeed, projects must comply with treaties as they must with other federal and state laws.

Courts consider any limitation on a tribe’s treaty fishing right sufficient grounds for halting the permitting process or denying permits altogether to government and private parties. For example, even though the government project in *Umatilla* included proposed mitigation efforts to trap and haul chinook salmon from below the proposed dam, the project could not proceed without express congressional action because access to a steelhead fishery would be eliminated. Likewise,
Private Party Liability Under Indian Treaties

in *Northwest Sea Farms v. United States Army Corps of Engineers*, the district court rejected the justification that a proposed private fish farm would have only a de minimis effect on the Lummi Tribe’s rights where the Indian fishers could still harvest fish at other stations. The site did not need to be the most primary or most productive; rather, the court reasoned that access to all usual and accustomed fishing stations was reserved under treaty. Similarly, in *Muckleshoot v. Hall*, though the Tribes could continue to catch the same fair portion of fish at stations outside of the proposed private project area, the district court denied an injunction against the tribe based, in part, on evidence that the Tribes would have to expend more money and time to catch “the same number of fish.”

Because treaties operate as an independent source of federal law, a tribe’s fishing right can serve as the sole ground on which a federal agency may deny a project permit. For example, in *Northwest Sea Farms*, the district court upheld a federal agency’s determination that denied a permit to the fish farm on the basis that it would impede the Lummi Tribe’s treaty-reserved right to access historic fishing stations. The court specifically refused to defer to an earlier state administrative proceeding that, if binding on the district court, would unilaterally extinguish a treaty right by means of a permitting process, rather than through congressional enactment.

Adherence to other federal or state laws does not indicate that a project complies with relevant Indian treaties. For example, the

---

108. Id. at 1522.
109. Id. at 1521.
111. See *Northwest Sea Farms*, 931 F. Supp. at 1522.
112. Id.
113. Id. at 1523 n.8. It should also be noted that statutory causes of action are not dispositive of treaty claims unless a comprehensive statute “speaks directly” to the question of remedies for treaty right impairment. See *County of Oneida v. Oneida Indian Nation*, 470 U.S. 226, 237 (1985) (citing *Milwaukee v. Illinois*, 451 U.S. 304, 315 (1981)). Likewise, where Congress does not authorize an administrative agency to award monetary compensation for past or present injury to a tribe, the fact that an agency reviews the impact of a contested project does not preempt a tribe’s claim for monetary relief in federal court. See *Nez Perce Tribe v. Idaho Power Co.*, 847 F. Supp. 791, 800–01 (D. Idaho 1994) (“FERC [Federal Energy Regulatory Commission] did not, and indeed could not, order monetary compensation for past or present injury to the fish runs. . . . Such an action is properly brought in the courts, not before FERC.”).
federal district court in *Muckleshoot* enforced the Muckleshoot and Suquamish Tribes’ treaty rights, even though the proposed private marina contractors conducted an extensive environmental review, procured long-sought-after federal and local permits, and would face significant financial harm. The Tribes sued the City of Seattle, the Army Corps of Engineers (the Corps), and a private developer to enjoin construction of a marina sited atop an historic fishing station just north of the already densely developed Seattle waterfront. The Corps estimated the Tribes’ financial losses at between $9335 and $40,000; the Tribes calculated the potential loss to Indian fishers as over $255,000 annually, which accounted for impacts of the marina itself. Resting its decision to grant injunctive relief solely on the possibility of irreparable injury to the treaty right, the court did not reach the federal statutory claims alleged under the National Environmental Policy Act. The treaty fishing right alone provided grounds to enjoin the private project.

Similarly, in *No Oilport! v. Carter*, the district court granted the Tribes’ request for an evidentiary hearing to determine whether a private company’s proposed oil pipeline would proximately cause a decline in the “size or quality” of the fish run. Granting summary judgment for the defendant on all other statutory environmental claims, the court characterized the Tribes’ claims as “the most troublesome of all the issues.”

**B. Federal Courts Award Monetary Relief Against the Federal Government for Interference with the Treaty Fishing Right**

In *Menominee Tribe of Indians v. United States*, the U.S. Supreme Court recognized that a treaty reserves to tribes a fishing interest equivalent to a bona fide property right. In doing so, the Court held

---

116. Id. at 1505–06.
117. Id. at 1506.
118. Id. at 1517.
119. Id.
121. Id. at 372.
122. Id. at 371.
124. Id. at 413.
Private Party Liability Under Indian Treaties

that the Menominees’ treaty entitled the Tribe to just compensation for any unlawful taking by the federal government. The Tribe’s right survived as a separate and cognizable property right despite assimilationist legislation by Congress that had previously terminated the Tribe’s official status. That is, even though Congress extinguished the federal trust supervision of tribal property and services, the Court refused to imply that the legislation likewise vanquished fishing and hunting rights.

Tribes have also secured monetary relief before the Indian Claims Commission, which Congress established to vindicate Indian rights via a waiver of U.S. sovereign immunity. As in the proceedings leading up to and affirmed by Menominee, for example, the Commission has awarded compensation for abrogation of treaty fishing rights. Short of full abrogation, the Commission also awarded monetary relief for the partial limitation of a treaty fishing right in Confederated Tribes of the Colville Reservation v. United States. In Colville, the Tribes brought suit against the United States for authorizing dams and commercial operations that depleted the on-reservation supply of fish. The Commission granted compensation for the retail value of fish to which the Tribe was entitled, less the value of fish actually received.

C. Federal Courts Award Monetary Relief Against Any Party that Interferes with Other Rights Reserved Under Treaty

As with the treaty fishing right, tribes are entitled to just compensation where the United States takes tribal property interests in

125. Id.
126. Id. at 411–13.
127. Id.
129. Menominee, 391 U.S. at 413 (1968), aff’d 388 F.2d 998 (Ct. Cl. 1967).
130. 43 Indian Cl. Comm’n 505, 525 (1978).
131. Id.
132. Id. To be sure, if such judgments operate as a one-time buy out, tribes today would be reluctant to assert similar claims in federal court. See State Dep’t of Ecology v. Yakima Reservation Irrigation Dist., 121 Wash. 2d 257, 291, 850 P.2d 1306, 1325 (1993) (holding Indian Claims Commission final judgment barred the Yakima Indians from subsequently protecting their treaty fishing rights under the doctrine of res judicata).
land that are reserved under treaty or recognized by statute. The federal government may not take a tribe’s treaty-reserved land by appropriating title to third parties without payment to the tribe as if the tribe owned the land in fee simple. Compensation is likewise due for federal taking of timber or mineral rights secured through the possessory rights inherent in treaty-reserved land.

Federal courts also award monetary relief where state or local governments interfere with treaty rights in land. For example, in County of Oneida v. Oneida Indian Nation, the Oneida Tribes sued the New York counties of Oneida and Madison for damages, alleging interference with their possessory right to occupy the area inhabited by the county citizenry. The U.S. Supreme Court sustained the Tribe’s common law trespass claim—although it arose 175 years prior—as a live federal issue and awarded damages to the Oneidas for the unlawful possession by the Counties. Against this backdrop, in United States v. Pend Oreille Public Utility District No. 1, the United States Court of Appeals for the Ninth Circuit sustained federal question jurisdiction and remanded the case to determine the appropriate injunctive and

134. Tribes acquire property interests by treaty, aboriginal possession, executive order, congressional establishment of Indian reservation, and other mechanisms. COHEN’S HANDBOOK, supra note 128, at 471–86. Underlying these forms of property interests is the assumption that, by virtue of discovery, the federal government holds land title in trust for tribes and thus retains legal “ownership”; however, a tribe’s right to “use” and “occupy” the land is exclusive and enforceable. Id. at 523–28. For property guaranteed by treaty, tribes are entitled to just compensation for its appropriation. See Creek Nation, 295 U.S. at 111.
135. Creek Nation, 295 U.S. at 110 (stating that the power of the United States to control and manage “did not enable the United States to give the tribal lands to others, or to appropriate them to its own purposes, without rendering, or assuming an obligation to render, just compensation for them; for that would not be an exercise of guardianship, but an act of confiscation”) (internal citation omitted).
138. Id. at 229.
139. Id. at 241 (“We think the borrowing of a state limitations period in these cases would be inconsistent with federal policy. Indeed, on a number of occasions Congress has made this clear with respect to Indian land claims.”).
140. Id. at 230.
141. 28 F.3d 1544 (9th Cir. 1994).
142. Id. at 1549 n.8 (noting that federal jurisdiction is not disputed). Though such disputes would typically be relegated to state courts, federal jurisdiction is sustained as arising under an Indian treaty and thus is within “the exclusive province of federal law.” See Oneida, 470 U.S. at 234–36.
monetary relief for the public utility’s flooding of the Kalispel Tribe’s land, which constituted a trespass.143

Likewise, tribes have a federal common law cause of action for damages to protect real property interests from private party interference.144 As early as 1850, in Marsh v. Brooks,145 the U.S. Supreme Court assumed an action for ejectment and remanded a case for trial based on the issue of interference with Indian possessory rights in land.146 Over a century later, two cases from the United States Court of Appeals for the Tenth Circuit recognized actions for damages to protect tribal lands from private parties.147 In Mescalero Apache Tribe v. Burgett Floral Co.,148 the court sustained claims for ejectment and recovery of monetary relief against a private party that harvested timber on reservation land.149 Similarly, in Pueblo of Isleta v. Universal Constructors, Inc.,150 the Tenth Circuit considered a treaty-based damages claim for injury to property resulting from nearby blasting activities of a private company.151 Rejecting the private defendant’s argument that tribes could not base a claim on land title held in trust by another,152 the Pueblo court reasoned that it was “not appropriate to bring into play subtle principles of English common law” to overlay a treaty right which stands uniquely apart from such constraints.153 Instead, as the Pueblo of Isleta court summarized, tribes are entitled to damages against a private party given the difficulty for individual tribal members to assert their rights in court, the property interests held in common by the Tribe, and the strong interest of the United States to ensure that the Tribe and its members receive “even-handed justice.”154

143. Pend Oreille, 28 F.3d at 1549–52.
145. 49 U.S. 223 (1850).
146. Id. at 232 (citing Johnson v. McIntosh, 21 U.S. 543 (1823)).
147. See Pueblo of Isleta v. Universal Constructors, Inc., 570 F.2d 300, 301–02 (10th Cir. 1978); Mescalero Apache Tribe v. Burgett Floral Co., 503 F.2d 336, 338 (10th Cir. 1974).
148. 503 F.2d 336 (10th Cir. 1974).
149. Id. at 338.
150. 570 F.2d 300 (10th Cir. 1978).
151. Id. at 302–03.
152. Id. at 302 (“The United States is actually the title owner.”).
153. Id. at 301 (citing Oneida Indian Nation v. County of Oneida, 414 U.S. 661, 677 (1974)); see also Mescalero, 503 F.2d at 338 (citing Oneida to protect a broad set of property interests under treaty in federal court, which included an action for damages against a trespasser).
154. Pueblo of Isleta, 570 F.2d at 302–03.
D. Courts Award Monetary Relief to Commercial Fishers at Common Law

The case law awarding monetary relief to tribes for interference with certain treaty rights is consistent with the remedies secured by fishers at common law. Although ownership over wildlife does not arise until the creature is reduced to capture, a fisher need not “own” fish in order to assert a compensable legal interest therein. For example, in *Columbia River Fishermen v. City of St. Helens*, the Oregon State Supreme Court awarded damages to commercial fishermen where a town and a paper mill polluted river waters, which interfered with the commercial catch. The claim was not one based on ownership of the fish, but rather involved a claim to protect the right of fishermen to pursue their vocation. In turn, this common-law right of fishers imposes a corollary duty on others to avoid imperiling fish populations; such interference presents a cause of action for damages in trespass, negligence, or nuisance.

In short, federal courts grant injunctive relief to halt government and private projects that, if constructed, would limit a tribe’s treaty fishing right. As an independent source of federal law, courts require treaty compliance notwithstanding a project’s conformity with state and other federal laws. In addition to injunctive relief, federal courts award monetary relief to tribes where the federal government interferes with a

155. See *Columbia River Fishermen v. City of St. Helens*, 87 P.2d 195, 197–98 (Or. 1939); Bales v. City of Tacoma, 172 Wash. 494, 498–504, 20 P.2d 860, 863–64 (1933); see also Sanders, *supra* note 42, at 166 nn.79–83 (discussing the history of cases illustrating the common law cause of action for damages for injury caused by interruption or interference with a person’s fishing rights). It should be noted, however, that harm to the treaty right is distinct from harm to commercial fishers at common law where the right to fish is held by the Tribe on behalf of its members, not individuals. See *Whitefoot v. United States*, 293 F.2d 658, 661–63 (Ct. Cl. 1961) (holding that a $15 million payment to Tribe for abrogation of treaty fishing rights included compensation of individual Indians).


157. See *St. Helens*, 87 P.2d at 197–98; see also *Geer*, 161 U.S. at 529 (affirming the state “ownership” doctrine to regulate wildlife within borders).

158. 87 P.2d 195 (Or. 1939).

159. *Id.* at 196–97.

160. *Id.*


162. See *supra* Part II.A.

163. See *supra* notes 114–122 and accompanying text.
treaty fishing right. Likewise, federal courts award monetary relief for interference with other rights reserved under treaty, regardless of the defendant actor. Finally, commercial fishers may secure damages at common law for interference with their vocational rights.

III. IN THE NINTH CIRCUIT, COURTS HAVE REJECTED TRIBAL FISH DAMAGE CLAIMS AGAINST PRIVATE PARTIES

With reservations in the shadow of hydroelectric dams, the Nez Perce and Skokomish Tribes each brought damage claims in federal court for past harms to their respective treaty fishing rights, which resulted from dam operations. Under the Federal Power Act (FPA), Congress exempts the United States from liability for any harm that results from dam operation and construction, leaving power companies alone to compensate for downstream harms. The Ninth Circuit granted a rehearing en banc in Skokomish Indian Tribe v. United States in February 2004 to consider the question of whether tribes have a cause of action for damages in trespass where a private company has interfered with their treaty right to fish.

A. Nez Perce Tribe v. Idaho Power Co.

In the first case of its kind, the U.S. District Court for the District of Idaho in Nez Perce Tribe v. Idaho Power Co. found jurisdiction to hear a tribe’s treaty claim for damages for the reduction in number of fish and access to customary stations caused by a private company’s

164. See supra Part II.B.
165. See supra Part II.C.
166. See supra note 155 and accompanying text.
168. See 16 U.S.C. § 803(c) (2000). Whether the federal government, in the first instance, can delegate away its trust obligation to tribes is beyond the scope of this Comment.
169. See Skokomish, 161 F. Supp. 2d at 1183 (dismissing claim against a federal defendant).
170. 161 F. Supp. 2d 1178 (W.D. Wash. 2001), aff’d in part, vacated in part by 332 F.3d 551 (9th Cir. 2003), and vacated by 358 F.3d 1180 (9th Cir. 2004) (granting rehearing en banc).
171. Id. As of this writing, the Ninth Circuit has not yet published its en banc decision.
dam construction and operation.\textsuperscript{173} Adopting the magistrate judge’s report to stand in for its own, the court reasoned that the FPA, under which the federal government licensed the dam, did not preempt the treaty-based claim where the available administrative process could not grant damage awards, but could only impose future mitigation measures on the dam license.\textsuperscript{174} The court denied that the Tribe has a right to preservation of fish runs as they stood at the 1855 treaty signing.\textsuperscript{175} Thus, the court reasoned that the Tribe has no modern-day right on which to base a claim for monetary relief.\textsuperscript{176} The tribe, however, prayed for monetary relief based on harm caused by construction and operation of the dam since 1955.\textsuperscript{177} Despite acknowledging the protection afforded to fishing rights in \textit{Winans}, \textit{Umatilla}, and \textit{Fishing Vessel}, the \textit{Nez Perce} court nevertheless denied damages on the grounds that injunctive relief awarded in prior cases could be distinguished from the monetary relief sought by the Nez Perce tribe.\textsuperscript{178} The court, via the magistrate, cited no authority for this proposition.\textsuperscript{179} Notably, monetary relief is the default remedy at common law; injunctive relief is granted only where the plaintiff shows that damages are inadequate at law.\textsuperscript{180}

Notwithstanding the \textit{Oneida} Court’s expansive reading of a tribal cause of action, the \textit{Nez Perce} court reasoned that the Tribe could not sustain a claim for damages based on its treaty fishing right.\textsuperscript{181} The court reasoned that the right to take fish was not plainly a “property interest” because the Tribe lacked ownership over “the fish runs themselves.”\textsuperscript{182}

\begin{footnotesize}
\begin{enumerate}
\item Id. at 794, 799. The court stated that federal question jurisdiction existed, as “it is beyond any reasonable dispute that the Tribe’s fishing rights and their claims in this regard are derived from the 1855 treaty.” Id. at 799 (citing 28 U.S.C. § 1362 (1988); Oneida Indian Nation v. County of Oneida, 414 U.S. 661, 676–78 (1974)). Indeed, federal jurisdiction over tribal claims is expressly provided for in 28 U.S.C. § 1362. See Moe v. Confederated Salish & Kootenai Tribes of Flathead Reservation, 425 U.S. 463, 472 (1976) (stating that the act was intended “to open the federal courts to the kind of claims that could have been brought by the United States as trustee, but for whatever reason, were not so brought”).
\item \textit{Nez Perce}, 847 F. Supp. at 803.
\item Id. at 807.
\item Id. at 807–13.
\item Id. at 794, 812.
\item Id. at 809.
\item See id. at 806–10.
\item See, e.g., Knaebel v. Heiner, 663 P.2d 551, 553 (Alaska 1983) (holding that injunctive relief is proper where damages are inadequate) (citing Coffman v. Breeze Corps., 323 U.S. 316, 322 (1945)).
\item Id.
\end{enumerate}
\end{footnotesize}
Private Party Liability Under Indian Treaties

In a lengthy footnote, the court noted that if the Tribe had a property interest in the fish, the Tribe would have a cause of action against "any private party who intentionally or negligently injured the fish." Because the U.S. Supreme Court in *Menominee* held that, indeed, fishing rights are property rights entitled to just compensation, the *Nez Perce* court then necessarily distinguished its holding, which denied monetary relief. The district court distinguished compensation for action taken by Congress to "deprive" a tribe of its fishing rights from a "reduction" of that same resource caused by a dam. Having thus characterized the case, the court applied institutional capacity arguments and cited disfavor for judicial activism to support its refusal to grant a "new" common law cause of action. Ultimately, the Nez Perce Tribe secured a multi-million dollar settlement from Idaho Power.

B. *Skokomish Indian Tribe v. United States*

Although the courts sustained jurisdiction in *Nez Perce* and *Oneida* for claims arising under treaty, the U.S. District Court for the Western District of Washington dismissed all treaty-based claims in *Skokomish Indian Tribe v. United States*. As the Nez Perce Tribe did, the Skokomish Tribe sued for damages to its treaty fishing right caused by an upstream hydroelectric project, which for eighty years had nearly eliminated the stream flow both on- and off-reservation. The district court dismissed the pleaded claims arising under treaty, after concluding that the claims sounded in state common law or arose from the Tribe’s objection to the facility license. The FPA, however, dictates the contours of the dam license and carries no private cause of action. Considering state common law claims, the court reasoned that harm to the fish runs did not pose a continuing injury, but a permanent and long-

183. Id. at 810 n.22.
184. Id. at 811 (citing Menominee Tribe of Indians v. United States, 391 U.S. 404, 413 (1968)).
185. Id.
186. Id. at 815.
190. Skokomish, 161 F. Supp. 2d at 1179.
191. Id. at 1179–80.
standing one; thus, the statutes of limitations foreclosed any remaining
claims.193 The court did not discuss the continuing harm to the treaty
right itself as defeating a state time bar.194

IV. FEDERAL COURTS SHOULD AWARD DAMAGES WHERE
PRIVATE PARTIES IMPAIR A TREATY FISHING RIGHT

The treaty right to take a fair portion of available fish, together with
the right of access, ensures that tribes have actual fish to harvest.195 The
use of injunctions to protect the treaty right before projects destroy fish
runs indicates that monetary relief is necessary to make the tribe whole
after the harmful action is taken.196 Such a remedy is consistent with the
U.S. Supreme Court’s characterization of the treaty fishing right as a
compensable property interest,197 with precedent that awards
compensation for interference with other treaty rights,198 and with
common law claims available to commercial fishers.199

A. The Right to Take Fish from Historic Fishing Stations Assumes
that There Are Actual Fish to Take.

Only in-the-flesh, actual fish can fulfill a tribe’s reserved right to take
a fair portion of fish.200 The Fishing Vessel and Winans Courts both
rejected attempts to reduce the fishing right to abstraction in the face of

193. Skokomish, 161 F. Supp. 2d at 1181–83. In contrast, in Mille Lacs Band of Chippewa
Indians v. Minnesota, 853 F. Supp. 1118, 1124–25 (D. Minn. 1994), aff’d, 124 F.3d 904 (8th Cir.
1997), and aff’d, 526 U.S. 172 (1999), the court declined to borrow a state statute of limitations
where the Band asserted a claim for injunctive and declaratory relief directly under a treaty fishing
right. Id. Instead, because the State of Minnesota continued to enforce natural resource regulations
against the Band, a wrong to their fishing and hunting rights was likely “continuing,” and thus the
limitations period for the treaty claim had not expired. Id. (citing 28 U.S.C. § 1262 (1988), which
conferred original jurisdiction over Indian treaty claims).


(1905); see also Northwest Res. Info. Ctr. v. Northwest Power Planning Council, 35 F.3d 1371,
1376 n.6 (9th Cir. 1994) (“The Court [in Fishing Vessel] also noted that the treaty guarantee of ‘the
right of taking fish’ was meaningful only if fish were available for the taking.” (emphasis in
original)).

196. See infra notes 213–216 and accompanying text.


198. See supra Part II.C.

199. See Columbia River Fishermen v. City of St. Helens, 87 P.2d 195, 196 (Or. 1939).

200. See supra Part I.B.
Private Party Liability Under Indian Treaties

competition from nontreaty fishers. The rights of access and fair apportionment work in tandem to ensure that tribal fishermen have actual fish to harvest in order to sustain a moderate livelihood. Guaranteeing access to fishing stations is the means to ensure the end of securing steady supplies of fish. In both Fishing Vessel and Winans, the Court contemplated actual numbers of fish for the tribal taking. Although the Fishing Vessel Court declined to specify a bare number of fish beyond a fifty-percent ceiling, presumably, where the Tribal harvest was at the time two percent of the total, the numbers spoke for themselves. The court reserved the future possibility of adjusting the percentage if needed to protect ceremonial and subsistence values, which further supports the conclusion that the treaties guarantee the right to take actual fish.

On the shoulders of Winans, federal courts should enforce the full extent of fishing rights—regardless of the defendant actor. Indeed, whether a treaty fishing right binds any particular party relates to the scope of the right. Federal courts are bound by U.S. Supreme Court precedent to broadly interpret the extent of the right, according to canons of construction which uniquely favor Tribal interests. Since the Winans decision a century ago, private parties are unquestionably subject to Indian treaties. More recently, federal courts have held that no party is authorized to limit a treaty fishing right without congressional authorization. The district courts in Sea Farms,
Muckleshoot, No Oilport!, and Umatilla uniformly granted injunctive or declaratory relief to protect treaty fishing rights against government agencies, private actors, and their regulatory counterparts.212 Courts should construe the right to access and take actual fish as equivalent to a tribe’s full, fair portion of the harvest taken from historic fishing grounds prior to project construction.213 As the Umatilla, Muckleshoot, and Sea Farms decisions indicate, the fact that a project allows a tribe to take some of the available salmon or to access some of the fishing stations does not satisfy the treaty right; rather, injunctive relief is proper for any limitation on the right.214 Courts also have enjoined projects that require more money and time for a tribe to catch even “the same number of fish,”215 or projects such as the private pipeline in No Oilport! that affect the “size or quality of the run.”216 Most notably, the Muckleshoot court considered the Tribe’s estimated financial loss should the private project proceed.217 Even though the opposing parties had different dollar estimates,218 the fact that the court considered the Tribe’s monetary losses indicates that courts consider a tribe’s right to the pre-project fair harvest of fish.219


213. See Umatilla, 440 F. Supp. at 555. Courts would likely consider only the loss attributed to the proposed project. See Muckleshoot, 698 F. Supp. at 1506 (citing dollar estimates of the potential loss of harvestable fish). This is consistent with the steps to bring “fish damage claims” outlined in a United States Department of the Interior memorandum. Memorandum from the Associate Solicitor, Indian Affairs, to Solicitor 1 (May 25, 1982). The memo outlined potential claims that the Department considered bringing on behalf of tribes against hydroelectric dam operators, elements of which included: (a) reasonably specific proof of the fishery prior to construction of the dam; (b) proof that the dam caused the loss of fish; and (c) a determination of the loss suffered. Id. The United States declined to pursue any such claims. Id. at 1, 30.

214. See Sea Farms, 931 F. Supp. at 1522; Muckleshoot, 698 F. Supp. at 1515; Umatilla, 440 F. Supp. at 555–56; see also United States v. Washington, 157 F.3d 630, 650 (9th Cir. 1998) (reversing trial court decision where the lower court had improperly limited the treaty scope in determining the remedy).


216. No Oilport!, 520 F. Supp. at 372 (quoting United States v. Washington, 506 F. Supp. 187, 208 (W.D. Wash. 1980), aff’d in part, rev’d in part, 694 F.2d 1374 (9th Cir. 1982)). To be sure, the No Oilport! court relied in part on the portion of the opinion in United States v. Washington later vacated by the Ninth Circuit; however, the No Oilport! court first cited Fishing Vessel to support the conclusion that a project threatening a tribe’s moderate standard of living must be adjudicated to ensure that a tribe’s fishing right is not limited by a unilateral private action. Id.


218. Id.

219. See Confederated Tribes of the Colville Reservation v. United States, 43 Indian Cl. Comm’n
Private Party Liability Under Indian Treaties

B. Monetary Relief Is Required to Make the Tribe Whole when Any Party Interferes with the Right to Take Actual Fish

Given that both government and private parties are subject to treaties, courts that impose monetary penalties against one defendant actor and not the other unlawfully limit the scope of the treaty right. Despite U.S. Supreme Court precedent in *Menominee* and the Indian Claims Commission’s decision in *Colville*, which construe the treaty fishing right as compensable property, the court in *Nez Perce* distinguished the elimination of tribal harvest by the federal government from that by a private party’s project and denied damages. Rather than discussing the extent of the legal right harmed, the *Nez Perce* court expressed concern over imposing a remedy that threatened to return Idaho to nineteenth-century conditions. The tribe, however, did not claim a right to catch pre-industrialization levels of fish, but rather prayed for monetary relief for the decline in fish runs, which was proximately caused by construction and operation of the dam since 1955. Not only did the *Nez Perce* court mischaracterize the claim, it also restricted the treaty right in prescribing the appropriate remedy. Courts can, at best, adjust the magnitude of the award as justice requires, but they are not at liberty to limit the scope of the right itself. To do otherwise is reversible error because Congress alone has the authority to alter treaty terms. Assuredly, the tribe would still have to prove that the project proximately caused the fish decline. See Memorandum from the Associate Solicitor, Indian Affairs, to Solicitor 1 (May 25, 1982).

505, 541 (1978). Note that the number of fish prior to project implementation is merely a characterization of the right, not necessarily the scope of the harm caused by a defendant project. Assuredly, the tribe would still have to prove that the project proximately caused the fish decline. See Memorandum from the Associate Solicitor, Indian Affairs, to Solicitor 1 (May 25, 1982).

220. *See supra* Part I.C.


224. *Id.* at 794.

225. *See id.* at 811–12 (discussing the scope of the right under the heading, “Award of Monetary Damages”).

226. *See United States v. Washington*, 157 F.3d 630, 650 (9th Cir. 1998); *Cree v. Waterbury*, 78 F.3d 1400, 1405 (9th Cir. 1996).

227. *See Washington*, 157 F.3d at 650; *Cree*, 78 F.3d at 1405.

228. *See supra* Part I.C.
from traditional common law remedies, Congress has yet to say so. 229 Indeed, Congress has affirmatively opened federal courthouse doors to tribes asserting claims arising under treaty. 230

The emphasis on the extent of the legal right harmed rather than the identity of the defendant actor is consistent with the federal judiciary’s across-the-board grant of monetary relief for interference with other treaty rights. 231 For treaty-guaranteed land, federal courts have sustained federal question jurisdiction and applied federal common law causes of action to compensate tribes for interference with tribal possessory rights at the hands of both private and government parties. 232 Damages are the appropriate remedy either as one-time compensation for Congress’s full abrogation of the treaty right, 233 or for harm proximately caused by a private defendant, as with blasting activities in Pueblo of Isleta, 234 or for a county’s unlawful possession, as in Oneida. 235 Awarding damages to tribes for private party interference with fishing rights, as the Court did in Menominee for federal party interference, is consistent with the rights-focused rationale in the treaty land context. 236 In addition, given the essential role of fishing rights to treaty negotiations, 237 courts should consider fish claims arising under treaty as analogous to land claims, which arise under the very same legal instrument. 238 Even if, as in Nez Perce, a court fails to construe the fishing right as property, courts

229. Courts have long sustained common law causes of action for treaty interference. See supra notes 144–145 and accompanying text. Congress has yet to expressly legislate to limit such causes of action.

230. 28 U.S.C. § 1362 (2000); see also Moe v. Confederated Salish & Kootenai Tribes of Flathead Reservation, 425 U.S. 463, 472 (1976) (stating the act was intended “to open the federal courts to the kind of claims that could have been brought by the United States as trustee, but for whatever reason, were not so brought”).

231. See supra Part II.B.

232. See supra Part II.C.


236. See supra Part II.B (discussing cases awarding monetary relief against both government and private parties).

237. See supra Part I.A.

238. See Mille Lacs Band of Chippewa Indians v. Minnesota, 853 F. Supp. 1118, 1124–25 (D. Minn. 1994) (borrowing the rationale used in the Oneida land title case as relevant in the fishing rights context) (citing Oneida, 470 U.S. at 240), aff’d, 124 F.3d 904 (8th Cir. 1997), and aff’d, 526 U.S. 172 (1999).
Private Party Liability Under Indian Treaties

should compensate tribal fishers’ like other non-Indian fishers. Indeed, under a common law claim by fishers, it is not necessary that tribes even “own” fish to assert a right upon which relief may be granted. The St. Helens court based remedies to fishers solely on the right of vocation, regardless of ownership over the resource. Additionally, as in St. Helens, remedies may attach to the fishers’ asserted rights not only for “a current supply of salmon,” but also for the “future supply of salmon” diminished by the private action. Remedies available to tribes should be, at the very least, equivalent to those available to commercial fishers, particularly given that tribes’ fishing rights are secured in enforceable, written legal instruments and contain rights broader than those tied to vocation. Thus, for example, even if the citizen suit provisions available under the Clean Water Act and Endangered Species Act preempt a St. Helens cause of action, federal treaties remain the supreme law of the land unless expressly altered by Congress.

V. CONCLUSION

In Winans, the U.S. Supreme Court warned that nothing less than the word of the nation stands behind a tribe’s reserved treaty right to fish. Tribes ceded millions of acres of land, contingent upon continued access to actual fish as necessary to support a livelihood, common sustenance, and ceremonial values. It is not enough, then, that private projects

239. Columbia River Fishermen v. City of St. Helens, 87 P.2d 195, 196 (Or. 1939); Sanders, supra note 42, at 164.
240. See St. Helens, 87 P.2d at 196; Sanders, supra note 42, at 164; see also Oneida Indian Nation v. County of Oneida, 414 U.S. 661, 677 (1974) (rejecting the argument that the Tribes must base their claim of possession on actual title); Pueblo of Isleta v. Universal Constructors, Inc., 570 F.2d 300, 302 (10th Cir. 1978) (rejecting defendant’s arguments based on Anglo-American private property constructs).
242. Id. at 196, 199 (noting the past harm to fishing rights as well as the thousands of dollars that would be lost in the future).
243. See Sanders, supra note 42, at 164 (arguing that the Nez Perce court’s reasoning not only makes the Tribe’s rights inferior to commercial fishers, but would preclude even common law claims).
244. See supra Part I.A.
245. See supra note 113 (discussing preemption by treaty).
246. See supra notes 101, 114–122 and accompanying text.
operate under a compendium of state and federal regulations. Without express congressional authorization, treaties that long-precede such permitting schemes require a private party’s compliance. Given that federal courts have enforced the treaty fishing right against private and government parties alike, courts should award damages against private actors just as they grant monetary relief for claims of past harms by government parties.

249. Winans, 198 U.S. at 382.
MAKING MOMMIES: THE WASHINGTON STATE COURT OF APPEALS EXCEEDED ITS AUTHORITY BY CREATING A COMMON LAW PARENTEAGE ACTION IN IN RE PARENTEAGE OF L.B.

Thomas G. Robinson-O’Neill

Abstract: In In re Parentage of L.B., Division I of the Washington State Court of Appeals created a new common law cause of action that allows a same-sex de facto parent to be declared a legal parent. In the alternative, the court held that a de facto parent has a cause of action under Washington’s nonparental visitation statute. This Note argues that the court exceeded its authority in creating a common law cause of action because the Uniform Parentage Act and the statutory scheme preempt the common law. Further, this Note argues that the ability of a de facto parent to obtain a visitation order under the nonparental visitation statute adequately protects the best interests of the child. Therefore, the Washington State Supreme Court should reverse Division I’s creation of a common law parentage action and affirm Division I’s alternative holding.

Courts today must address an increasingly familiar family framework: a lesbian couple who has a child via artificial insemination. In most cases, one of the women is the child’s biological, and therefore legal, parent; courts recognize her as the child’s parent, and the legal system protects her relationship with the child. The other woman may be able to demonstrate that she is a “de facto parent” who acts in every way as a parent.


2. See Jacobs, supra note 1, at 342 (citing generally Julie Shapiro, De Facto Parents and the Unfulfilled Promise of the New ALI Principles, 35 WILLAMETTE L. REV. 769 (1999)).

3. For purposes of clarity, this Note will use the term “de facto parent” to describe a nonparent with a psychological relationship with a child equivalent to that of a biological parent. See In re Marriage of Allen, 28 Wash. App. 637, 648, 626 P.2d 16, 23 (1981). Case law from Washington and other jurisdictions addresses these people also as “psychological parents” and as those acting in loco parentis. See, e.g., In re Dependency of J.H., 117 Wash. 2d 460, 469, 815 P.2d 1380, 1384 (1991) (“In certain limited circumstances our appellate courts have recognized the importance of psychological parents to the welfare of children.”); In re Custody of S.H.B., 118 Wash. App. 71, 79, 74 P.3d 674, 680 (2003) (“Nonparents—including adoptive parents, legal guardians, grandparents, and persons acting in an in loco parentis capacity—do not have the same constitutional rights of a parent, absent legislative action.”). See generally Miller, supra note 1 (citing cases addressing de facto parents, psychological parents, and those acting in loco parentis). The various terms used by different jurisdictions refer to those who have a significant psychological relationship with children.
the child’s parent but remains a legal stranger.\(^4\) However, if the women’s relationship dissolves, the distinction between legal and de facto parents means that the law protects only the legal parent’s relationship with the child.\(^5\)

In Washington, until *In re Parentage of L.B.*\(^6\), a de facto parent had limited legal recourse to maintain a relationship with the child.\(^7\) However, in *In re Parentage of L.B.*, Division I of the Washington State Court of Appeals created a common law parentage action that allows a nonbiological de facto parent standing to be declared a legal parent when he or she can show that (1) the nonbiological parent consented to the formation of a parent-like relationship; (2) the nonbiological parent lived in the same household with the child; (3) the nonbiological parent assumed the obligations of parenthood without expectation of financial compensation; and (4) the nonbiological parent had been in a parental role long enough to form a bonded psychological relationship with the child.\(^8\)

This Note argues that Division I of the Washington State Court of Appeals (Division I) erred in recognizing a new common law parentage action. Although Washington courts have the authority to create common law, statutes and judicial caution limit this authority.\(^9\) Because the legislature intended the Uniform Parentage Act (UPA)\(^10\) to serve as the exclusive means to determine parentage and because there are other statutory remedies for de facto parents, Division I exceeded its authority in creating a common law cause of action.\(^11\) Part I describes how, prior

---

4. See Jacobs, supra note 1, at 342 (noting that the de facto parent had “no legal parental role because of her lack of biological connection to the child”).
5. See id. at 347 (noting that the law often fails to recognize a lesbian coparent’s parental status).
7. See, e.g., id. at 470–76, 89 P.3d at 276–79 (discussing statutory remedies for same-sex de facto parents); see also infra Part I.
8. L.B., 121 Wash. App. at 476, 89 P.3d at 279.
Making Mommies

to *In re Parentage of L.B.*, same-sex de facto parents could not be considered legal parents under the UPA. Part II addresses the remedies de facto parents had prior to Division I’s decision. Part III describes judicial authority to create common law. Part IV presents the court’s reasoning in *In re Parentage of L.B.* Part V argues that Division I exceeded its authority in creating a common law cause of action. This Note concludes that Division I should have held that de facto parents may obtain visitation orders under the nonparental visitation statute but should not be declared legal parents under the common law.

I. UNDER THE UPA, THE SAME-SEX PARTNER OF A BIOLOGICAL PARENT IS NOT A LEGAL PARENT

Courts do not typically consider the same-sex partner of a biological parent to be the child’s legal parent.\(^\text{12}\) At common law, legal parents were traditionally the biological mother and father.\(^\text{13}\) Before the advent of genetic testing, the difficulty of conclusively establishing biological relationships between fathers and children led to a common law presumption that spouses were the biological, and therefore the legal, parents.\(^\text{14}\) Because same-sex couples cannot marry in Washington,\(^\text{15}\) those who have children via artificial insemination do not fit neatly into this framework, as typically one of them is a biological parent and the other is not.\(^\text{16}\) The nonbiological or de facto parent may establish parentage by a statutory second parent adoption\(^\text{17}\) or an adjudication of parentage under the UPA.\(^\text{18}\)

The same-sex partner of a biological parent may pursue a second parent adoption.\(^\text{19}\) Adoption proceedings are purely statutory\(^\text{20}\) and allow

\(^\text{12}\) See Jacobs, *supra* note 1, at 343–44.


\(^\text{15}\) *Wash. Rev. Code* § 26.04.020(1) (2002) (“Marriages . . . are prohibited: . . . [w]hen the parties are persons other than a male and a female.”). Two separate Washington State Superior Courts recently declared this statute unconstitutional, but the decisions are stayed pending review by the Washington State Supreme Court. Tracy Johnson, *2nd Judge Rules Against State’s Gay Marriage Ban*, SEATTLE POST-INTELLIGENCER, Sept. 8, 2004, at A1. If the decision is upheld, a lesbian married couple who conceives via artificial insemination will be covered by the UPA.

\(^\text{16}\) See Jacobs, *supra* note 1, at 342.


\(^\text{19}\) *D.R.M.*, 109 Wash. App. at 189, 34 P.3d at 891 (citing *Wash. Rev. Code* § 26.33). This Note focuses on post-dissolution remedies for a de facto parent, because in most of these situations a
a court to declare that a person not biologically related to the child is a legal parent. The adoption statute is silent on the subject of second parent adoptions, but the Washington State Supreme Court has held that the standing requirements for adoption are very broad. Division I of the Washington State Court of Appeals has indicated that same-sex second parent adoptions are not prohibited under the adoption statutes.

Washington’s UPA authorizes actions to establish parentage, but it does not provide a de facto same-sex parent with standing. The UPA is intended to govern all parentage actions in Washington and gives courts the power to adjudicate parentage, including parenting responsibilities. The statute specifically allows for an adjudication of maternity, and it indicates that the provisions allowing for an legal parent is unlikely to consent to a second parent adoption.

23. See, e.g., In re Adoption of B.T., 150 Wash. 2d 409, 417–19, 78 P.3d 634, 638–39 (2003) (holding that the standing doctrine is very broad and allowing grandparents to adopt after intervening in a dependency hearing, as long as requirements of WASH. REV. CODE § 26.23.140 are met).
25. WASH. REV. CODE § 26.26.021 (“This chapter governs every determination of parentage in this state.”).
26. See id. §§ 26.26.021–.913. The UPA recognizes seven categories of persons with standing: (1) the child; (2) the mother; (3) a man whose paternity is to be adjudicated; (4) the division of child support; (5) an authorized adoption agency; (6) a representative authorized by law to act for a deceased, incapacitated, or minor individual; or (7) an intended parent under a surrogate parentage contract. See id. § 26.26.505.
27. Id. § 26.26.021(1). This statute acknowledges that the UPA does not interfere with rights and duties under other Washington laws. Id. § 26.26.021(3). Presumably, this refers to the dependency, adoption, and marital dissolution statutes, which all provide for determinations of the scope of parental rights. See WASH. REV. CODE §§ 13.34.010–.810 (2002) (dependency statute); WASH. REV. CODE §§ 26.09.002–.914 (marital dissolution statute); id. §§ 26.33.010–.910 (adoption statutes). While the UPA savings clause preserves the scope of parental authority under these statutes, it does not indicate that parentage may be determined by other means. Id. § 26.26.021(3). The standing provisions of the UPA, therefore, take into account other statutory methods of being declared parents. For example, the standing provisions of the UPA clearly denote that adoptive parents may be considered parents under the UPA. Id. § 26.26.101(1)(c), (2)(d). Therefore, this savings clause should not be read to preserve subsequently created common law action that creates legal parents rather than merely defining the scope of those rights.
29. Id. § 26.26.130(7) (providing for allocation of parenting responsibilities pursuant to WASH. REV. CODE § 26.09, the dissolution of marriage statute).
30. Id. § 26.26.101(1)(b) (stating that a mother-child relationship is established by an adjudication of maternity).
adjudication of paternity may also be used to determine a mother-child relationship.31

The UPA initially contained language that might have supported same-sex parentage actions.32 A former version of the statute allowed a man to establish parentage if, for example, he held the child out as his own or acknowledged paternity in writing.33 Division I noted in dicta that these provisions might apply to a same-sex couple through the provision that allows paternity actions to be used to determine maternity.34 However, in 2002, the legislature repealed this section of the UPA.35

A nonbiological partner in a same-sex relationship cannot use the current UPA to adjudicate parenthood.36 There is no provision in the UPA that provides for a determination of parentage on the basis of a person’s significant relationship with a child.37 In the case of artificial insemination, the UPA affords the spouse of the woman who conceived a presumption of fatherhood.38 Because same-sex couples cannot marry in Washington and the UPA specifies that only a husband enjoys such a presumption, a lesbian de facto parent cannot rely on this provision.39 Nor do the provisions concerning surrogate contracts 40 allow a lesbian de facto parent to be deemed an intended parent served by a “surrogate father” because the statute is gender specific.41

---

31 Id. § 26.26.051 ("[T]he provisions relating to determination of paternity may be applied to a determination of maternity.").


36 L.B., 121 Wash. App. at 474–75, 89 P.3d at 278 (concluding that the current UPA does not provide a cause of action for a same-sex unmarried individual to pursue parentage).


38 Id. § 26.26.710.

39 See, e.g., L.B., 121 Wash. App. at 474, 89 P.3d at 278 (concluding that the current UPA does not provide a cause of action for a same-sex unmarried individual to pursue parentage).


41 See L.B., 121 Wash. App. at 474, 89 P.3d at 278 (noting that the UPA’s surrogacy provisions are gender specific).
II. PRIOR TO IN RE PARENTAGE OF L.B., LEGAL PARENTS HAD RIGHTS SUPERIOR TO DE FACTO PARENTS

A de facto parent does not have the same rights that a legal parent has under the United States Constitution. Under Washington law, de facto parents have limited rights, but these statutory nonparental rights are circumscribed by the legal parents’ constitutional rights.

A. Legal Parents Have a Constitutional Right to Protect Their Relationships with Their Children, but that Right Is Not Absolute

The U.S. Supreme Court has repeatedly reaffirmed that parents have a fundamental liberty interest in the care, custody, and control of their children. Under the Fourteenth Amendment to the U.S. Constitution, a legal parent has a constitutional right to direct his or her child’s upbringing. These rights rest on a rebuttable presumption that a parent acts in the best interests of his or her child; thus, for instance, a parent may restrict the child’s contact with certain people because it is assumed

44. See id. at 20, 969 P.2d at 30.
45. See In re Custody of Nunn, 103 Wash. App. 871, 884, 14 P.2d 175, 182 (2001) (“Between a parent and a nonparent, therefore, a more stringent balancing test is required to justify awarding custody to the nonparent.”).
46. See, e.g., Troxel, 530 U.S. at 65 (plurality opinion) (noting “extensive precedent . . . that the Due Process Clause of the Fourteenth Amendment protects the fundamental right of parents to make decisions concerning the care, custody, and control of their children”); Washington v. Glucksberg, 521 U.S. 702, 720 (1997) (citing Meyer v. Nebraska, 262 U.S. 390, 399–400 (1923), for the proposition that it is a fundamental Due Process right to direct the education of one’s children); Santosky v. Kramer, 455 U.S. 745, 753 (1982) (describing the “fundamental liberty interest” of natural parents); Parham v. J.R., 442 U.S. 584, 602 (1979) (noting parents have a right to raise their children); Quillon v. Walcott, 434 U.S. 246, 255 (1978) (“[T]he relationship between the parent and child is constitutionally protected . . . by the Due Process Clause of the Fourteenth Amendment.”); Wisconsin v. Yoder, 406 U.S. 205, 232–33 (1972) (noting it is well-established that parents have a fundamental liberty interest in their children); Stanley v. Illinois, 405 U.S. 645, 651 (1972) (noting the integrity of the family unit is protected under the Due Process Clause); Pierce v. Soc’y of Sisters, 268 U.S. 510, 534–35 (1925) (indicating that a parent’s right to control the education of a child derives from the Fourteenth Amendment).
47. See U.S. CONST. amend. XIV, § 1 (“[N]or shall any State deprive any person of life, liberty, or property without due process of law.”); Meyer, 262 U.S. at 399–400 (recognizing that “liberty” in the Fourteenth Amendment includes “establish[ing] a home and bring[ing] up children”).
the parent acts in the child’s best interest. ⁴⁹ A court must defer to the decisions of a legal parent and may not override these decisions absent a compelling state interest. ⁵⁰ Only a legal parent has these constitutionally protected rights; nonparents, even de facto parents, do not have constitutionally protected interests in their relationships with their children. ⁵¹

Though a parent has constitutional rights, these rights are not absolute. ⁵² The state’s parens patriae ⁵³ authority allows the state to intervene in family decisions to protect the welfare of its citizens. ⁵⁴ When a fundamental liberty interest of a parent is involved, a state may intervene only where it has a compelling interest. ⁵⁵ The U.S. Supreme Court has upheld a state’s authority to impinge on a legal parent’s autonomy, or even completely sever a parent’s rights, where a child has been harmed or where there is a threat of harm to a child. ⁵⁶

Under Washington’s nonparental custody statute, the state’s power to override a legal parent’s wishes can be used to protect a de facto parent’s

⁴⁹. Smith, 137 Wash. 2d at 21, 969 P.2d at 31.
⁵⁰. See Troxel, 530 U.S. at 68–69 (plurality opinion). What constitutes a compelling state interest is not settled. The Washington State Supreme Court has held that only prevention of harm to the child constitutes such a compelling state interest. Smith, 137 Wash. 2d at 18, 969 P.2d at 29. On appeal, the Troxel Court declined to reach this question. See Troxel, 530 U.S. at 73 (plurality opinion).
⁵¹. Miller v. California, 355 F.3d 1172, 1176 (9th Cir. 2004) (finding that de facto parents have a right to appear in dependency proceedings, but have no other right or weightier constitutional interest); In re Custody of S.H.B., 118 Wash. App. 71, 80, 74 P.3d 674, 680 (2003) (holding that absent legislative action, nonparents, even those acting as in loco parentis, lack constitutional rights of parents).
⁵². Smith, 137 Wash. 2d at 16–17, 969 P.2d at 28–29.
⁵⁴. Smith, 137 Wash. 2d at 15–16, 969 P.2d at 28.
⁵⁵. Id. (citing, inter alia, Roe v. Wade, 410 U.S. 113, 155 (1973)). The Washington State Supreme Court has gone further and held that the remedy must serve a state’s compelling interest. Id. The U.S. Supreme Court, however, has not reached the question of whether the remedy must be narrowly tailored when it conflicts with a legal parent’s constitutional rights. See Sally F. Goldfarb, Visitation for Nonparents After Troxel v. Granville: Where Should States Draw the Line?, 52 RUTGERS L.J. 783, 785–86 (2001) (arguing that the Court failed to agree on the proper standard of review).
⁵⁶. See Wisconsin v. Yoder, 406 U.S. 205, 230–32 (1972). In its most recent case addressing the subject, however, the U.S. Supreme Court refused to consider whether a showing of harm was necessary where a case involved visitation by nonparents. See Troxel v. Granville, 530 U.S. 57, 73 (2000) (plurality opinion).
relationship with a child.\textsuperscript{57} In order to receive custody under this statute, a petitioner must either show that the legal parent does not have physical custody or allege that neither legal parent is a suitable custodian.\textsuperscript{58} A legal parent is not a suitable custodian when he or she causes actual detriment to the child.\textsuperscript{59} To show actual detriment, a petitioner must do more than convince the court that it is not in the child’s best interests to remain with the legal parent, but need not show parental unfitness.\textsuperscript{60}

The visitation provisions in Washington’s nonparental custody statute are even broader than its custodial provisions.\textsuperscript{61} The statute allows “[a]ny person” to petition for visitation “at any time,” as long as the visitation is in the best interests of the child.\textsuperscript{62} In 1998, the Washington State Supreme Court held that the statute was so broad as to be facially unconstitutional under the Due Process Clause of the U.S. Constitution because it infringed on fundamental rights of parents and did not serve a compelling state interest.\textsuperscript{63} In \textit{Troxel v. Granville},\textsuperscript{64} the U.S. Supreme Court affirmed the Washington State Supreme Court’s decision, but held the statute unconstitutional only as applied because the trial court did not

\begin{itemize}
\item \textsuperscript{57} WASH. REV. CODE §§ 26.10.010–.913 (2002); see also \textit{In re Custody of Stell}, 56 Wash. App. 356, 365, 783 P.2d 615, 620 (1989) (concluding that a de facto parent can be awarded custody under WASH. REV. CODE § 26.10 (1988)).
\item \textsuperscript{58} WASH. REV. CODE § 26.10.030(1). If the child is in the custody of a parent, the statute requires an allegation that neither legal parent is a suitable parent. \textit{Id.} Division I has interpreted the statute to require that the petitioner prevail on this element if the child is in the custody of a parent and the petitioner wishes to gain custody. \textit{In re Custody of S.H.B.}, 118 Wash. App. 71, 79, 74 P.3d 674, 679 (2003).
\item \textsuperscript{59} See, e.g., \textit{In re Marriage of Allen}, 28 Wash. App. 637, 649, 626 P.2d 16, 23 (1981) (“In extraordinary circumstances, where placing the child with an otherwise fit parent would be detrimental to the child, the parent’s right to custody is outweighed by the state’s interest in the child’s welfare.”).
\item \textsuperscript{60} \textit{Id.} ("There must be a showing of actual detriment to the child, something greater than the comparative and balancing analyses of the ‘best interests of the child’ test . . . [b]ut unfitness of the parent need not be shown."); see also \textit{In re Custody of Shields}, 120 Wash. App. 108, 111, 84 P.3d 905, 906 (2004) (holding that the nonparental custody action requires a showing of parental unfitness or actual detriment); \textit{In re Custody of R.R.B.}, 108 Wash. App. 602, 615–16, 31 P.3d 1212, 1219–20 (2001) (affirming an award of custody to the biological parent where the trial court found placement with the otherwise fit legal adoptive parents would be detrimental to the child); \textit{Stell}, 56 Wash. App. at 365, 783 P.2d at 620 (affirming an award of custody to an aunt when placement with the father would cause actual detriment to the child).
\item \textsuperscript{61} See \textit{Troxel}, 530 U.S. at 67 (plurality opinion) (“The Washington nonparental visitation statute is breathtakingly broad.”).
\item \textsuperscript{62} WASH. REV. CODE § 26.10.160(3).
\item \textsuperscript{63} See \textit{In re Custody of Smith}, 137 Wash. 2d 1, 19, 969 P.2d 21, 30 (1998), aff’d on other grounds sub nom. \textit{Troxel v. Granville}, 530 U.S. 57 (2000) (plurality opinion).
\item \textsuperscript{64} 530 U.S. 57 (2000).
\end{itemize}
Making Mommies

give deference to the legal parent’s wishes.65 Division I has subsequently held that the visitation provision remains valid as long as it complies with the limitations imposed by Troxel.66 Because deference to a parent’s wishes may be overcome where a child is harmed,67 and Washington courts recognize that severing a de facto parent-child bond could cause harm,68 a visitation order under the provisions of the nonparental visitation statute could comply with Troxel.69

B. Prior to In re Parentage of L.B., Washington Courts Recognized the Importance of De Facto Parents in Limited Circumstances

Recognizing the importance of the psychological relationship between children and de facto parents, Washington courts have awarded visitation or custody to de facto parents in limited circumstances.70 The

65. See id. at 69, 73, 75 (plurality opinion) (holding that the trial court violated Granville’s parental rights by not giving special weight to her wishes). The U.S. Supreme Court did not reach the question of whether the Due Process Clause requires a showing of harm to satisfy the compelling state interest condition precedent to granting visitation. Id. at 73.

66. See In re Parentage of L.B., 121 Wash. App. 460, 490, 89 P.3d 271, 286 (2004) (“[T]he third-party visitation statute is facially valid when construed in accord with the principles of Troxel . . . .”). Whether this is a correct analysis of the effect of Troxel on the Washington State Supreme Court’s holding in Smith is beyond the scope of this Note. At the very least, it remains an open question.

67. See In re Custody of R.R.B., 108 Wash. App. 602, 616, 31 P.3d 1212, 1219–20 (2001); see also Smith, 137 Wash. 2d at 15–16, 969 P.2d at 28 (“[T]he state may step in and override a decision of a parent where the decision would harm the child.”).

68. See, e.g., Smith, 137 Wash. 2d at 20, 969 P.2d at 30 (noting that where a significant relationship with a child is severed, there is potential for harm to the child).

69. See, e.g., L.B., 121 Wash. App. at 490, 89 P.3d at 286 (holding that the third party visitation statute could be constitutionally applied in the case of a de facto lesbian parent); see also In re Parentage of C.A.M.A., 120 Wash. App. 199, 214, 84 P.3d 1253, 1260 (2004) (“Finally, Troxel also required special factors which might justify the state’s interference with a parent’s fundamental right to make decisions. We conclude that one such factor is a longstanding de facto parental relationship between the petitioner and the child.”).

70. See, e.g., In re Dependency of J.H., 117 Wash. 2d 460, 469, 815 P.2d 1380, 1384 (1991) (“In certain limited circumstances our appellate courts have recognized the importance of psychological parents to the welfare of children.”); McDaniels v. Carlson, 108 Wash. 2d 299, 313, 738 P.2d 254, 263 (1987) (remanding a dispute between a putative and natural father with the direction that both be given the opportunity to petition for reasonable visitation rights); In re Custody of Shields, 120 Wash. App. 108, 130–31, 84 P.3d 905, 916 (2004) (affirming custody award to stepmother after the death of the child’s father because the child had been well-integrated into the family of the stepparent); In re Dependency of Ramquist, 52 Wash. App. 854, 862, 765 P.2d 30, 35 (1988) (affirming an award of custody to foster parents after the parent’s rights were terminated because the child was psychologically bonded to the foster family); In re Marriage of Allen, 28 Wash. App. 637, 649, 626 P.2d 16, 23 (1981) (awarding custody to the stepmother upon determination that placement with the father would cause actual detriment to the child).
courts have recognized that harm to the child may occur when a significant relationship between the child and an adult is severed. In *In re Custody of Smith*, the Washington State Supreme Court noted that “[i]n certain circumstances where a child has enjoyed a substantial relationship with a third person, arbitrarily depriving the child of the relationship could cause severe psychological harm to the child.”

Washington courts have recognized the importance of the stability of these relationships even when the adult involved is not a legal or natural parent.

Prior to *L.B.*, the rights of those de facto parents who could demonstrate such a significant relationship were typically limited to situations in which the rights of the legal parents did not conflict with those of the de facto parent. For example, the Washington State Supreme Court held that an aunt and uncle who had commenced a nonparental custody action and been awarded temporary custody had a right to intervene in a concurrent dependency action to terminate the legal parents’ rights. Thus, in a dependency action in which the legal parent’s rights are severed, a de facto parent could intervene and obtain custody. Courts have awarded custody or visitation to nonparents with psychological bonds in the context of stepparent divorces and contested paternity actions.

---

71. See, e.g., Smith, 137 Wash. 2d at 20, 969 P.2d at 30 (recognizing potential for harm where a substantial relationship is severed); J.H., 117 Wash. 2d at 469, 815 P.2d at 1384 (acknowledging the importance of relationships between children and nonparent caretakers); McDaniels, 108 Wash. 2d at 310–11, 738 P.2d at 261–62 (recognizing a bond between child and two putative fathers); Allen, 28 Wash. App. at 649, 626 P.2d at 23 (recognizing the bond between a child and stepmother).


73. Id. at 20, 969 P.2d at 30.

74. “Child development experts widely stress the importance of stability and predictability in parent/child relationships, even where the parent figure is not the natural parent.” McDaniels, 108 Wash. 2d at 310, 738 P.2d at 261 (citing generally, inter alia, J. Goldstein, et al., Beyond the Best Interests of the Child (1979)); see Allen, 28 Wash. App. at 648, 626 P.2d at 23 (“Where the reason for deferring to parental rights—the goal of preserving families—would be ill-served by maintaining parental custody, as where a child is integrated into the nonparent’s family, the de facto family relationship does not exist as to the natural parent and need not be supported.”).

75. See, e.g., J.H., 117 Wash. 2d at 471, 815 P.2d at 1385 (reasoning that foster parents have a permissive right to intervene so long as their intervention does not conflict with the legal parent’s rights).


78. See McDaniels v. Carlson, 108 Wash. 2d 299, 313, 738 P.2d 254, 263 (1987) (holding that
C. Prior to In Re Parentage of L.B., a De Facto Parent’s Remedies Were Limited

Although a court could award to a de facto parent either visitation or custody over the wishes of a legal parent, a de facto parent has fewer remedies and rights.79 In Smith v. Organization of Foster Families,80 the U.S. Supreme Court addressed foster parents’ contention that the U.S. Constitution protected their psychological parent relationships with the children in their care.81 The Court assumed, without deciding, that a right to protect a substantial relationship with a child did exist and held that New York’s procedural protections for foster parents were sufficient.82 The Court cautioned that constitutional rights for nonparents such as foster parents would necessarily be in tension with a legal parent’s constitutional rights.83 Therefore, the Court concluded that where a foster parent’s wishes conflict with a legal parent’s wishes, the rights of the foster parent “must be substantially attenuated where the proposed removal from the foster family is to return the child to his natural parents.”84

Though the U.S. Supreme Court has not explained when rights are substantially attenuated,85 Washington courts make this determination by carefully balancing the rights of the child’s nonparents, legal parents, and the state.86 In In re Custody of Nunn,87 Division I of the Court of

---

79. In re Custody of S.H.B., 118 Wash. App. 71, 80, 74 P.3d 674, 680 (2003) (holding that absent legislative action, nonparents, even those acting as in loco parentis, lack constitutional rights of a parent); Miller v. California, 355 F.3d 1172, 1176 (9th Cir. 2004) (finding that de facto parents have a right to appear in dependency proceedings, but have no other right or weightier interest of constitutional dimension).
81. See id. at 846–47.
82. Id. at 847.
83. See id. at 846; accord Troxel v. Granville, 530 U.S. 57, 64 (2000) (plurality opinion) (“The extension of statutory rights in this area to persons other than a child’s parents, however, comes with an obvious cost. For example, [it] can place a substantial burden on the traditional parent-child relationship.”).
84. See Org. of Foster Families, 431 U.S. at 847.
85. See Troxel, 530 U.S. at 78 (Souter, J., concurring) (“Our cases, it is true, have not set out exact metes and bounds to the protected interest of a parent in the relationship with his child . . . .”).
86. See, e.g., In re Custody of Smith, 137 Wash. 2d 1, 19 n.4, 969 P.2d 21, 30 n.4 (1998)
Appeals reversed the trial court’s award of custody to a paternal aunt over a fit biological mother after the death of the biological father.\textsuperscript{88} The court reasoned that a stringent balancing test is required in a custody dispute between a legal parent and a nonparent to justify awarding custody to a nonparent.\textsuperscript{89} Thus, the court held that in order to minimize state interference with a family unit, a nonparent must produce evidence of parental unfitness before a court can award custody to the nonparent.\textsuperscript{90} Unless the nonparent can show the child is being or will be harmed, a nonparent has no standing to challenge a legal parent’s rights.\textsuperscript{91}

Even where a de facto parent’s asserted rights did not conflict with a legal parent’s,\textsuperscript{92} Washington courts have acknowledged that the nonparent had limited due process protection in comparison to the legal parent.\textsuperscript{93} At common law, nonparents, even grandparents, had no legal right to petition for visitation or custody.\textsuperscript{94} Similarly, a nonparent is not entitled to the same presumptions as a legal parent within the (acknowledging that the court had applied a balancing analysis and affirming the result because the threshold showing of harm was met), \textit{aff’d on other grounds sub nom. Troxel,} 530 U.S. 57 (plurality opinion); \textit{id.} at 24, 969 P.2d at 32 (Talmadge, J., concurring in part, dissenting in part) (noting that the court must balance the interests of the child, legal parent, and state before awarding visitation); \textit{In re Custody of Nunn,} 103 Wash. App. 871, 884, 14 P.3d 175, 182 (2001) (reasoning that a more stringent balancing test is required when determining whether to award custody or visitation to a parent or a nonparent); \textit{In re Custody of Anderson,} 77 Wash. App. 261, 264, 890 P.2d 525, 526 (1995) (holding that in a custody dispute between parents and nonparents, the analysis must accommodate the natural parent’s constitutional interests.); \textit{In re Marriage of Allen,} 28 Wash. App. 637, 646, 626 P.2d 16, 22 (1981) (applying a balancing test among parental rights, state interests, and child’s interests to determine when a nonparent may be awarded custody).

\textsuperscript{87} 103 Wash. App. 871, 14 P.3d 175 (2001).

\textsuperscript{88} \textit{Id.} at 873–74, 14 P.3d at 177.

\textsuperscript{89} See \textit{id.} at 884, 14 P.3d at 182. The strongest expression of this approach comes from \textit{Smith,} in which the Washington State Supreme Court held that a petitioner must make a showing of harm to the child before the court can intervene to order visitation. \textit{Smith,} 137 Wash. 2d at 20, 969 P.2d at 30.

\textsuperscript{90} See \textit{Nunn,} 103 Wash. App. at 883, 14 P.3d at 181.

\textsuperscript{91} See \textit{id.}

\textsuperscript{92} In Washington, a foster parent may not be involved, for example, where the state has instituted a dependency hearing to terminate a legal parent’s rights. \textit{In re Dependency of J.H.,} 117 Wash. 2d 460, 471–72, 815 P.2d 1380, 1385 (1991) (reasoning that foster parents have a permissive right to intervene as long as their intervention does not conflict with the legal parent’s rights).


\textsuperscript{94} See Martin, \textit{supra} note 53, at 579.
Making Mommies

nonparental custody statute.\(^{95}\) In practice, this means that a court can override the nonparent’s wishes without providing the same due process protections granted to a parent.\(^{96}\) As a corollary, once a nonparent ceases to act as the child’s parent, the nonparent’s legal obligation to support the child ends.\(^{97}\)

Thus, before In re Parentage of L.B., there was a significant difference between the rights of de facto and legal parents.\(^{98}\) As a nonparent petitioning under the nonparent custody and visitation statute, a de facto parent could ask a court to intervene only when there is actual detriment to the child.\(^{99}\) Even then, the due process rights of the nonparent with a court order were limited in comparison to those of a legal parent.\(^{100}\)

III. STATUTES AND JUDICIAL CAUTION LIMIT THE AUTHORITY OF WASHINGTON COURTS TO CREATE COMMON LAW

Washington courts have the authority to create common law, but this authority is limited.\(^{101}\) A court may not, for example, contravene statutes or statutory intent.\(^{102}\) Moreover, the Washington State Supreme Court has indicated that a court should be cautious in creating common law “absent prior legislative or judicial expression on the subject.”\(^{103}\) In In re

\(^{95}\) See WASH. REV. CODE §§ 26.10.010–.913 (2002); S.H.B., 118 Wash. App. at 80–82, 74 P.3d at 680–81 (“Nonparents—including adoptive parents, legal guardians, grandparents, and persons acting in an in loco parentis capacity—do not have the same constitutional rights of a parent, absent legislative action.”).

\(^{96}\) See, e.g., Smith v. Org. of Foster Families, 431 U.S. 816, 853 (1977) (noting that the “balance of due process interests must accordingly be different” between a foster parent and a natural parent).


\(^{98}\) See, e.g., S.H.B., 118 Wash. App. at 80, 74 P.3d at 680 (noting that nonparents do not have the same constitutional rights as parents).

\(^{99}\) See supra notes 58–60 and accompanying text.

\(^{100}\) S.H.B., 118 Wash. App. at 80, 74 P.3d at 680.


\(^{102}\) See, e.g., id. (noting that the court may not rewrite unambiguous statutes to suit the court’s notion of public policy).

\(^{103}\) Thompson v. St. Regis Paper Co., 102 Wash. 2d 219, 232, 685 P.2d 1081, 1089 (1984) (citation omitted) (indicating that courts should be cautious in expanding the common law tort for wrongful discharge unless there is a prior judicially or legislatively recognized public policy favoring such expansion).
Parentage of L.B., Division I cited recent decisions to articulate that a court may create common law where there is a void in the statutes and a clear public policy mandate.\textsuperscript{104}

In Washington, courts have the authority and duty to alter the common law as justice requires.\textsuperscript{105} The common law consists of court-made rules,\textsuperscript{106} fashioned to adjust the law to deal with the “rich variety” of cases confronting the courts.\textsuperscript{107} The court’s authority over the common law may extend to creating new judicial remedies.\textsuperscript{108} In \textit{Ueland v. Reynolds Metals Co.},\textsuperscript{109} for example, the Washington State Supreme Court, recognizing a child’s independent cause of action for loss of parental consortium, noted that to defer to the legislature would “abdicate [the court’s] responsibility to reform the common law to meet the evolving standards of justice.”\textsuperscript{110}

A court’s authority over the common law is limited by statute.\textsuperscript{111} The common law gives way to statutory authority when statute and common law conflict.\textsuperscript{112} Courts are guided by statutory intent when they determine whether a statute abrogates the common law.\textsuperscript{113} The goal of statutory interpretation is to give effect to the intent and purpose of the legislature.\textsuperscript{114} The court may not substitute its own interpretation of public policy in place of the legislature’s.\textsuperscript{115} This prohibition includes

\textsuperscript{105} Lundgren v. Whitney’s, Inc., 94 Wash. 2d 91, 95, 614 P.2d 1272, 1275 (1980) (“Indeed, we have often discharged our duty to reassess the common law and alter it where justice requires.”).
\textsuperscript{106} Spokane Methodist Homes, Inc. v. Dep’t of Labor & Indus., 81 Wash. 2d 283, 286, 501 P.2d 589, 591 (1972).
\textsuperscript{107} Lundgren, 94 Wash. 2d at 95, 614 P.2d at 1275 (quoting Rodriguez v. Bethlehem Steel Corp., 525 P.2d 669, 676 (Cal. 1974)).
\textsuperscript{108} Ueland v. Reynolds Metals Co., 103 Wash. 2d 131, 136, 691 P.2d 190, 193 (1984) (“When justice requires, this court does not hesitate to expand the common law and recognize a cause of action.”).
\textsuperscript{109} 103 Wash. 2d 131, 691 P.2d 190 (1984).
\textsuperscript{110} Id. at 136, 691 P.2d at 193.
\textsuperscript{112} Id. at 222, 517 P.2d at 587.
\textsuperscript{113} In re Custody of Smith, 137 Wash. 2d 1, 8, 969 P.2d 21, 24–25 (1998) (“The purpose of statutory interpretation is to determine and give effect to legislative intent. Legislative intent is primarily determined from statutory language.” (citations omitted)), aff’d on other grounds sub nom. Troxel v. Granville, 530 U.S. 57 (2000) (plurality opinion).
\textsuperscript{114} Harmon v. Dep’t of Soc. & Health Servs., 134 Wash. 2d 523, 530, 951 P.2d 770, 773 (1998).
\textsuperscript{115} See State v. Jackson, 137 Wash. 2d 712, 725, 976 P.2d 1229, 1235 (1999) (refusing to read failure to meet common law duty of support to a child as a crime under accomplice liability statute).
“resist[ing] the temptation to rewrite an unambiguous statute to suit [the court’s] notions of what is good public policy . . .”\textsuperscript{116}

The court also may not adopt a common law cause of action that conflicts with an existing statutory scheme.\textsuperscript{117} Nor may the court create a cause of action where the legislature intends the statutory remedy to be exclusive.\textsuperscript{118} Thus, although the court may create common law in an area of law in which the legislature has acted,\textsuperscript{119} the court’s authority is strongest where the cause of action originated in the common law\textsuperscript{120} or “in a field peculiarly nonstatutory . . . .”\textsuperscript{121} In a field of law in which the legislature has acted, the court may act only where there is a void in the statutes.\textsuperscript{122}

In Philippides v. Bernard\textsuperscript{,123} the Washington State Supreme Court held that a court may not act on matters where the legislature has created a comprehensive set of statutes.\textsuperscript{124} The court refused to acknowledge a common law cause of action for loss of consortium on behalf of the parents of an adult dependent.\textsuperscript{125} A majority of eight justices held that because the legislature had created a comprehensive set of statues governing who can recover in wrongful death actions, the court was precluded from acting.\textsuperscript{126} Writing for the majority, Justice Faith Ireland observed that creating a common law remedy would directly conflict

\begin{footnotesize}
\begin{enumerate}
\item "resist[ing] the temptation to rewrite an unambiguous statute to suit [the court’s] notions of what is good public policy . . ."\textsuperscript{116}
\item Roth v. Bell, 24 Wash. App. 92, 100, 600 P.2d 602, 607 (1979) (“The legislature may legislate in areas dealing with facets of tort law involving familial relationships, but such legislative action does not divest courts of their authority to alter rules of law which had their genesis in the courts under the common law.”).
\item Stanard v. Bolin, 88 Wash. 2d 614, 617, 565 P.2d 94, 96 (1977) (“Because the action has its origins in the common law and has not been acted upon by the legislature, it is proper for us to reexamine it and determine its continued viability in light of present-day society.”).
\item Roberts, 92 Wash. App. at 655, 966 P.2d at 378–79 (determining that a court may use common law to fill such a void); \textit{see also In re Parentage of L.B.}, 121 Wash. App. 460, 476 n.2, 89 P.3d 271, 279 n.2 (2004) (indicating that a common law remedy survives the enactment of a statutory remedy if the common law fills a void).
\item Id. at 390, 88 P.3d at 946 (holding that the court may not create a common law beneficiary for wrongful death and survival actions).
\end{enumerate}
\end{footnotesize}
with existing statutes.\footnote{127}

In addition to heeding the limitations imposed by statutes, a court should proceed cautiously when creating a common law cause of action.\footnote{128} In developing a new common law cause of action for the tort of wrongful discharge,\footnote{129} the Washington State Supreme Court warned that the court “should not create public policy but instead recognize only clearly existing public policy . . . .”\footnote{130} The court indicated that the legislature is the fundamental source of public policy and that courts should not act upon previously unrecognized public policy.\footnote{131} However, the Washington State Supreme Court also indicated that in other circumstances\footnote{132} it would extend common law if there was a “compelling necessity.”\footnote{133} Thus, Washington courts should proceed cautiously when creating a new common law cause of action.\footnote{134}

IV. IN RE PARENTOAGE OF L.B., DIVISION I CREATED A COMMON LAW PARENTAGE ACTION

In In re Parentage of L.B., Division I of the Washington State Court of Appeals addressed Sue Ellen Carvin’s petition to be declared a parent of L.B. under Washington’s UPA.\footnote{135} Prior to L.B., there was a

\footnote{127. Id.}


\footnote{129. The court created this cause of action to ameliorate the common law rule of employment terminable at will in Thompson, 102 Wash. 2d at 232–33, 685 P.2d at 1089. The new cause of action allows a common law suit for wrongful discharge where there is a clear public policy against allowing the discharge. Id. The court adopted this exception to the common rule despite state legislative action against wrongful discharge. Id. at 226, 685 P.2d at 1086.}

\footnote{130. Sedlacek v. Hillis, 145 Wash. 2d 379, 390, 36 P.3d 1014, 1020 (2001).}

\footnote{131. Id. at 390, 36 P.3d at 1019.}

\footnote{132. Erhardt v. Havens, Inc., 53 Wash. 2d 103, 106–107, 330 P.2d 1010, 1012 (1958) (declining to allow a child to maintain an independent cause of action for loss of consortium where the father conceded he had a statutory remedy). Although Erhardt was never overruled, the Washington State Supreme Court later adopted this cause of action in Ueland v. Reynolds Metals Co., 103 Wash. 2d 131, 136, 691 P.2d 190, 193 (1984).}

\footnote{133. Erhardt, 53 Wash. 2d at 106, 330 P.2d at 1012; see also In re Parentage of L.B., 121 Wash. App. 460, 476 n.2, 89 P.3d 271, 279 n.2 (2004) (adopting the language from Thompson and Sedlacek to require a clear mandate of public policy before acting); Roth v. Bell, 24 Wash. App. 92, 100, 600 P.2d 602, 607 (1979) (interpreting Erhardt to require a compelling necessity before adopting a new cause of action).}

\footnote{134. See Thompson, 102 Wash. 2d at 232, 685 P.2d at 1089.}

\footnote{135. See L.B., 121 Wash. App. at 465, 89 P.3d at 274.}
significant difference between the rights of de facto and legal parents.\textsuperscript{136} The court in \textit{L.B.} held that Washington recognizes a common law parentage action and adopted a test previously articulated by the Wisconsin Supreme Court.\textsuperscript{137} The court’s decision, allowing de facto parents to be declared legal parents, gives de facto parents the constitutional protections that they did not previously have.\textsuperscript{138} The court also held in the alternative that a lesbian nonbiological parent had a cause of action under the visitation provisions of the nonparental custody statute.\textsuperscript{139}

Sue Ellen Carvin and Page Britain, a lesbian couple, became romantically involved in 1989 and began cohabiting.\textsuperscript{140} After five years, they decided Britain would have a child via artificial insemination, and they procured sperm from John Auseth, a gay male friend.\textsuperscript{141} Carvin and Britain acted as the child’s primary caretakers for the first six years of the child’s life.\textsuperscript{142} Although the couple discussed adoption, Carvin did not adopt L.B.\textsuperscript{143}

In 2001, when L.B. was six, Britain decided to separate from Carvin.\textsuperscript{144} Initially, the couple shared parenting of L.B., but shortly after the split Britain and Auseth terminated Carvin’s visitation with L.B.\textsuperscript{145} Carvin then filed a petition to be declared L.B.’s legal parent.\textsuperscript{146} The trial court found that L.B. had been harmed by the separation from Carvin, but held that Carvin had no cause of action under the UPA.\textsuperscript{147} The trial

\begin{itemize}
\item \textsuperscript{136} See, e.g., \textit{In re Custody of S.H.B.}, 118 Wash. App. 71, 80, 74 P.3d 674, 680 (2003) (noting nonparents do not have the same constitutional rights as parents).
\item \textsuperscript{137} \textit{L.B.}, 121 Wash. App. at 485, 487, 89 P.3d at 285.
\item \textsuperscript{138} \textit{Id.} at 485, 487, 89 P.3d at 285 (holding that the common law cause of action would allow a court to award a de facto parent shared parentage). In other words, their rights are the same after they are declared de facto parents. \textit{Id.}
\item \textsuperscript{139} \textit{Id.} at 488, 89 P.3d at 285.
\item \textsuperscript{140} \textit{Id.} at 465, 89 P.3d at 274.
\item \textsuperscript{141} \textit{Id.} at 465–66, 89 P.3d at 274.
\item \textsuperscript{142} \textit{Id.} at 467, 89 P.3d at 275.
\item \textsuperscript{143} \textit{Id.} at 467, 89 P.3d at 274.
\item \textsuperscript{144} \textit{Id.} at 467, 89 P.3d at 275.
\item \textsuperscript{145} \textit{Id.} at 467–68, 89 P.3d at 275.
\item \textsuperscript{146} \textit{Id.} at 468, 89 P.3d at 275. Shortly after Carvin filed her petition, Britain married Auseth, and he signed a declaration of paternity. \textit{Id.} Britain argued at the appellate court that Auseth should be allowed to intervene in the proceeding. \textit{Id.} at 490, 89 P.3d at 286. The court declined to reach the question of Auseth’s place in the proceeding, but remanded for the trial court to make an initial determination. \textit{Id.} Therefore, this Note will not address Auseth’s standing.
\item \textsuperscript{147} \textit{Id.} at 469, 89 P.3d at 276.
\end{itemize}
court also declined to create a common law parentage action and determined that the nonparental visitation statute did not apply because Carvin had not shown that Britain was an unfit parent.\(^{148}\)

On appeal, Division I reversed the trial court’s decision and recognized a common law cause of action that allows de facto parents to adjudicate parentage.\(^{149}\) The court held that Carvin did not have a cause of action under Washington’s UPA.\(^{150}\) Specifically, it held that although L.B. was conceived via artificial insemination, the artificial insemination provisions in the UPA applied only to married couples.\(^{151}\) Nor could Carvin rely on the surrogate parent standing provisions because the statute’s definition of “surrogate parent” was gender specific.\(^{152}\) The court determined that the UPA was unambiguous and, because the UPA did not cover Carvin’s situation, she had no cause of action.\(^{153}\)

Nonetheless, the court reasoned that “our Legislature cannot have intended to leave parties to such parentage arrangements and the children born of such arrangements without any remedy whatsoever.”\(^{154}\) The court then held that courts may create a common law remedy when the remedy “fills a void in the law for redress of an act or omission that contravenes a clear mandate of public policy.”\(^{155}\) Under this test, the court reasoned that Washington’s prior judicial recognition of de facto parentage provided the public policy mandate necessary to justify creating a common law cause of action.\(^{156}\)

To determine the parameters of the common law doctrine, the court turned to other jurisdictions that have recognized a common law cause

---

148. Id.
149. Id. at 487, 89 P.3d at 285.
151. Id. at 474, 89 P.3d at 278.
152. Id. at 473–74, 89 P.3d at 278; see also WASH. REV. CODE § 26.26.210(3)–210(4) (2002) (specifying that a surrogate is a female who gestates a child and that a surrogate contract means a contract in which a female gestates a child and relinquishes parental rights).
154. Id. at 475–76, 89 P.3d at 279.
155. Id. at 476 n.2, 89 P.3d at 279 n.2.
156. Id. The court also cited Washington’s Equal Rights Amendment as a source of public policy. Id. The court, however, did not address that issue in depth, id. at 476, 89 P.3d at 278–79, and it is beyond the scope of this Note.
Making Mommies

of action for de facto parents such as Carvin. Specifically, the court followed the test delineated by the Wisconsin State Supreme Court in In re Custody of H.S.H.-K. Division I held that Carvin must show on remand that (1) Britain consented to and fostered the parent-like relationship; (2) Carvin and L.B. lived together in the same household; (3) Carvin assumed the obligations of parenthood without expectation of financial compensation; and (4) Carvin had been in the parental role long enough to establish a dependent, bonded parental relationship.

Division I’s formulation, however, exceeded that of the Wisconsin court by extending the doctrine to full parentage rather than merely visitation. The Wisconsin court articulated a doctrine that allows a same-sex de facto parent standing to bring a visitation petition, but dismissed the custody action on the grounds that proving the existence of a parent-like relationship was insufficient to justify custody. Similarly, other jurisdictions that have adopted equitable parent doctrines have stopped short of allowing a de facto parent to be adjudged a legal parent.

In the alternative, the L.B. court held that a de facto parent had standing to petition the court for visitation under the Washington nonparental custody statute. The court adopted the actual detriment test for supporting an award of visitation. The court further implied that a de facto parent would likely meet this actual detriment standard by showing that the legal parent’s decision to stop visitation might harm the child. On this independent ground, the court indicated that Carvin may

157. Id. at 481–85, 89 P.3d at 282–83.
158. 533 N.W.2d 419 (Wis. 1995); L.B., 121 Wash. App. at 487, 89 P.3d at 285.
160. See id.
161. H.S.H.-K., 533 N.W.2d at 436 (discussing Wisconsin’s statutory basis for denying legal parents custody). The court also noted that the custody statutes in Wisconsin were intended to be the exclusive means of determining parentage. See id. at 431.
162. Id. at 420 (“We affirm that part of the order dismissing the petition for custody . . . .”).
163. See T.B. v. L.R.M., 786 A.2d 913, 920 (Pa. 2001) (affirming an award of partial custody and visitation); cf. V.C. v. M.J.B., 748 A.2d 539, 554 (N.J. 2001) (“Once a third party has been determined to be a psychological parent . . . he or she stands in parity with the legal parent.”). Despite this quote, the court did not award custody rights to the third party, only visitation rights. Id. (holding that visitation is the presumptive rule); see also Jacobs, supra note 1, at 366–69 (discussing the ways in which equitable doctrines in Wisconsin, New Jersey, and Massachusetts do not give full parental protections).
165. Id. at 489, 89 P.3d at 286.
be entitled to visitation rights.\textsuperscript{167}

V. DIVISION I EXCEEDED ITS AUTHORITY IN CREATING A COMMON LAW DE FACTO PARENTAGE ACTION

Division I of the Washington State Court of Appeals exceeded its authority when it created a common law parentage action for de facto parents. The Washington State Legislature intended Washington’s UPA to be the exclusive means of establishing paternity.\textsuperscript{168} Moreover, when the UPA is read in combination with the nonparental custody statute, the statutory scheme effectively preempts common law parentage actions.\textsuperscript{169} Even if the UPA and nonparental visitation statute were not exclusive, Division I’s own test for when courts can create a common law cause of action requires a clear mandate of public policy.\textsuperscript{170} The nonparental custody statute, which allows a de facto parent to petition for visitation, adequately prevents harm to the child while protecting the legal parent’s constitutional rights.\textsuperscript{171} Given the complex constitutional issues involved, the court should not have created a common law parentage action.

A. Division I Exceeded Its Authority in Creating a Common Law Cause of Action Because the UPA Is Intended to Be the Exclusive Means of Determining Parentage

Because the UPA was intended to be the exclusive means of determining parentage,\textsuperscript{172} a court may not create a common law parentage action. Washington courts do not hesitate to create common

\textsuperscript{166} See id. at 489, 89 P.3d at 285–86 (“Under Washington case law . . . Carvin does not need to prove that Britain is unfit in the classical sense, but only that it is detrimental to the child to sever the very parent-child relationship that Britain first consented to and fostered.”).

\textsuperscript{167} Id.


\textsuperscript{169} See, e.g., Philippides v. Bernard, 151 Wash. 2d 376, 390, 88 P.3d 939, 946 (2004) (explaining that a comprehensive statutory scheme leaves no room for a court to create common law).

\textsuperscript{170} See L.B., 121 Wash. App. at 476 n.2, 89 P.3d at 279 n.2.

\textsuperscript{171} Id. at 490, 89 P.3d at 286 (concluding that when properly applied, the nonparental custody statute recognizes the presumption that a fit parent will act in the best interests of the child and its remedy is narrowly tailored); see also In re Custody of H.S.H.-K., 533 N.W.2d 419, 436 (Wis. 1995) (explaining that de facto parent visitation protects parental constitutional rights and the child’s best interest).

\textsuperscript{172} See WASH. REV. CODE § 26.26.021.
Making Mommies

law where it is in the interest of justice.\footnote{See supra notes 105–110 and accompanying text.} The common law authority of the courts is, however, limited both by statute\footnote{See supra notes 111–118 and accompanying text.} and by the judicial canon of caution.\footnote{See supra notes 128–133 and accompanying text.} Division I recognized this in \textit{In re Parentage of L.B.} by noting that it could create a common law cause of action only where there was a statutory void and a clear mandate of public policy.\footnote{In re Parentage of L.B., 121 Wash. App. 460, 476 n.2, 89 P.3d 271, 279 n.2 (2004).} The court’s opinion, however, fails its own test.

The UPA does not grant standing to same-sex couples to adjudicate parentage.\footnote{See supra notes 36–41 and accompanying text.} The current version of the statute, as Division I held, “clearly does not” allow same-sex couples to pursue parentage under the UPA.\footnote{L.B., 121 Wash. App. at 474–75, 89 P.3d at 278.} The court further noted that the standing provisions of the UPA were unambiguous and that a same-sex partner like Carvin did not fit any of the unambiguous categories granted standing.\footnote{Id. at 472, 89 P.3d at 278.} Thus, the plain language of the statute does not allow same-sex de facto parents to be declared legal parents.

Moreover, the UPA is intended to be the exclusive method for determining parentage.\footnote{See WASH. REV. CODE § 26.26.021 (2002) (defining the scope of UPA).} The statute expressly states: “This chapter governs every determination of parentage in this state.”\footnote{Id. § 26.26.021(1).} The court acknowledged this plain language,\footnote{L.B., 121 Wash. App. at 476 n.2, 89 P.3d at 279 n.2.} but concluded that because the legislature intentionally excluded same-sex couples who have children via artificial insemination, the legislature intended for the omission to be remedied by the courts.\footnote{Id. at 475–76, 89 P.3d at 279.} This interpretation is inconsistent with the statute’s unambiguous language stating that the UPA governs every parentage determination in Washington.\footnote{See WASH. REV. CODE § 26.26.021.} The very existence of a common law parentage action directly contradicts the statute, and because a court should not “rewrite an unambiguous statute to suit [the court’s] notions of what is good public policy . . . ,”\footnote{State v. Jackson, 137 Wash. 2d 712, 725, 976 P.2d 1229, 1235 (1999).} Division I exceeded its authority in creating a common law parentage action,
regardless of its public policy rationale. Additionally, despite Division I’s conclusion that the legislature “cannot have intended” to leave same-sex couples without remedy,\(^\text{186}\) the omission of same-sex couples from the UPA should be interpreted to mean the legislature intended to exclude same-sex couples.\(^\text{187}\) The legislature amended the UPA in 2002 and repealed the clause that may have applied to same-sex couples.\(^\text{188}\) The decision to repeal those provisions should be read as intentional exclusion of same-sex couples from the UPA.

B. The Statutory Scheme Preempts Division I’s Creation of a Common Law Parentage Action

Division I’s conclusion that the legislature intended the courts to fill the void in the UPA with common law\(^\text{189}\) is further undermined because de facto parents have remedies under Washington’s adoption\(^\text{190}\) and nonparental custody statutes.\(^\text{191}\) The court’s own alternative holding that Carvin had a statutory remedy under the nonparental custody statute\(^\text{192}\) means that there is no statutory void to fill. A de facto parent can, for example, pursue a second parent adoption.\(^\text{193}\) Because an adoptive parent is a legal parent\(^\text{194}\) and there is no prohibition against same-sex adoptions,\(^\text{195}\) the adoption statute can provide a same-sex de facto parent with full constitutional protections.\(^\text{196}\)

\(^{186}\) L.B., 121 Wash. App. at 475–76, 89 P.3d at 279.
\(^{187}\) See Jackson, 137 Wash. 2d at 724–25, 976 P.2d at 1235 (admonishing the court to resist rewriting an unambiguous statute to suit the court’s public policy rationale); see also L.B., 121 Wash. App. at 473, 89 P.3d at 278 (noting that a lesbian nonbiological parent does not fit into any of the unambiguous categories for standing).
\(^{189}\) L.B., 121 Wash. App. at 476, 89 P.3d at 279.
\(^{190}\) See supra notes 17–22 and accompanying text.
\(^{192}\) L.B., 121 Wash. App. at 488, 89 P.3d at 285.
\(^{194}\) WASH. REV. CODE § 26.33.260 (2002); see D.R.M., 109 Wash. App. at 189, 34 P.3d at 891.
\(^{195}\) See D.R.M., 109 Wash. App. at 191, 34 P.3d at 892.
Making Mommies

Even after the dissolution of a same-sex relationship, Washington’s statutory scheme provides a remedy for a same-sex de facto parent via the nonparental statute.197 If the de facto parent can demonstrate that the child will suffer actual detriment, she can be awarded visitation or custody.198 Because Washington courts already recognize that severing a significant relationship, such as a de facto parent-child relationship, may cause children harm, a de facto parent could demonstrate this actual detriment.199 Division I acknowledged this in In re Parentage of L.B. by holding that a de facto parent had standing to petition for visitation under the nonparental statute.200

Thus, the statutory scheme preempts a common law remedy. A court cannot act in a field in which a comprehensive set of statutes exists.201 In Philippides, the court refused to extend the common law to allow parents to recover for loss of consortium caused by the death of an adult child.202 The court reasoned that the legislature had created a comprehensive wrongful death statute that defined statutory beneficiaries.203 To add parents of adult children to those beneficiaries would alter the legislative directive and exceed judicial authority.204 Similarly, with the UPA the legislature has created a comprehensive parentage statute and specified whom the statute covers.205 As in Philippides, creating a common law parentage action exceeds the court’s authority to modify legislative policy decisions.

197. See WASH. REV. CODE §§ 26.10.010–.913; see also In re Custody of Stell, 56 Wash. App. 356, 365, 783 P.2d 615, 620 (1989) (finding that a de facto parent can be awarded custody under WASH. REV. CODE § 26.10).


199. Smith, 137 Wash. 2d at 20, 969 P.2d at 30.


202. Id.

203. Id.

204. Id.

C. Lesbian De Facto Parents Have Appropriate Statutory Remedies and Division I Cannot Override the Legislature’s Policy Decision

Given the remedies available under the nonparental custody and visitation statute, the public policy arguments advanced by Division I are insufficient to justify a new cause of action. To justify such an action, the court would need a “clear mandate of public policy.” The difficulties in balancing the rights of legal parents, de facto parents, and children, however, obfuscate the public policy of creating a new parentage action. The actual detriment standard used by Washington courts to address petitions under the nonparental custody and visitation statute represents a careful balance between the rights of involved parties. As long as visitation orders are sufficient to remedy the identified risk of harm to the child, a court should not create an entirely new cause of action simply because the court believes it is better policy.

Because the court identified the prevention of harm to children as the relevant public policy, the court should have interpreted the nonparental custody and visitation statute as an adequate remedy for those parties, including same-sex couples, excluded from the UPA and marriage statutes. Using a visitation order to protect the de facto parental relationship is consistent with the Wisconsin court’s approach in *H.S.H.-K.*, the case on which Division I modeled its test. The Wisconsin court limited its de facto doctrine to visitation and deemed visitation sufficient to protect the child’s best interests, as well as parental autonomy and

---

209. *L.B.*, 121 Wash. App. at 481, 89 P.3d at 282; see also *In re Custody of R.R.B.*, 108 Wash. App. 602, 616, 31 P.3d 1212, 1219 (2001) (“[I]n a nonparental custody proceeding, the child’s safety, welfare, growth or development is always at issue; otherwise, there is no basis for awarding custody to a nonparent.”); *Smith*, 137 Wash. 2d at 15–16, 969 P.2d at 28 (“[T]he state may step in and override a decision of a parent where the decision would harm the child.”).
Making Mommies

constitutional rights.211 As the court noted in In re Parentage of L.B., a custody or visitation order under the existing nonparental statute allows a de facto parent to maintain a relationship with the child as long as the de facto parent can demonstrate that severing contact will cause detriment to the child.212 This approach carefully balances the legal parent’s rights with the child’s interests in maintaining a de facto parent relationship.213

VI. CONCLUSION

Division I of the Washington State Court of Appeals exceeded its authority in creating a common law parentage cause of action. Thus, the Washington State Supreme Court should reverse the Division I’s decision in In re Parentage of L.B. The UPA’s exclusivity clause preempts a court from creating a common law parentage cause of action. Even if the UPA were not the exclusive means of determining parentage, the UPA, adoption, and nonparental custody statutes create a statutory framework that preempts common law remedies. This does not, however, invalidate Division I’s concern with protecting the best interests of the child. The court should defer to the legislature’s policy decision to address this situation outside the context of the UPA. Thus, the Washington State Supreme Court should affirm Division I’s alternative holding that same-sex de facto parents have a remedy under the nonparental custody statute.

211. In re Custody of H.S.H.-K., 533 N.W.2d 419, 436 (Wis. 1995).
212. L.B., 121 Wash. App. at 489–90, 89 P.3d at 286.
213. See supra notes 85–91 and accompanying text.
AUTHOR INDEX

Alexander, Gerry  
Technology, Values, and the Justice System: Introduction 79:1

Andersen, William R.  
Technology and the Washington State Administrative Process—Some Preliminary Notes 79:13

Berry, Matthew R.  
Does Delaware’s Section 102(b)(7) Protect Reckless Directors from Personal Liability? Only if Delaware Courts Act in Good Faith 79:1125

Brust, John M.  
First Things First: Federal Courts Should Determine the Legal Status of a Lloyd’s of London Syndicate Before Deciding the Syndicate’s Citizenship for Diversity Purposes 79:969

Caughey, Michael  
Keeping Attorneys from Trashing Identities: Malpractice as Backstop Protection for Clients Under the United States Judicial Conference’s Policy on Electronic Court Records 79:407

Cerf, Vinton G.  
Internet and the Justice System 79:25

Cherry, Miriam A.  

Chiu, Anne Y.  
When Prisoners Are Weary and Their Religious Exercise Burdened, RLUIPA Provides Some Rest for Their Souls 79:999

Drobac, Jennifer Ann  
Sex and the Workplace: “Consenting” Adolescents and a Conflict of Laws 79:471

Gaul, Joshua C.  

Grogan, Kyla C.E.  
Lucky for Life: A More Realistic and Reasonable Estate Tax Valuation for Nontransferable Lottery Winnings 79:1153

Halm, Lindsay  
Putting Flesh on the Bones of United States v. Winans: Private Party Liability Under Treaties that Reserve Actual Fish for the Tribal Taking 79:1181

Harmon, Louise  
Wild Dreamers: Meditations on the Admissibility of Dream Talk 79:575

Himma, Kenneth Einar  
Towards a Theory of Legitimate Access: Morally Legitimate Authority and the Right of Citizens to Access the Civil Justice System 79:31

Horowitz, Donald J  

Horwitz, Morton J.  
Conceptualizing the Right of Access to Technology 79:105

Kaufman-Osborn, Timothy V.  
Capital Punishment, Proportionality Review, and Claims of Fairness (With Lessons from Washington State) 79:775

Meachum, Lynette  

Nissenbaum, Helen  
Privacy as Contextual Integrity 79:119
Perkins, David J.

Ramasastry, Anita
Government-to-Citizen Online Dispute Resolution: A Preliminary Inquiry 79:159

Robinson-O'Neill, Thomas G.

Seul, Jeffrey R.
Settling Significant Cases 79:881

Silverman, Gregory M.
Rise of the Machines: Justice Information Systems and the Question of Public Access to Court Records over the Internet 79:175

Small, T.W.
Boiko, Robert
Zorza, Richard
Designing an Accessible, Technology-Driven Justice System: An Exercise in Testing the Access to Justice Technology Bill of Rights 79:223

Sweeney, Dennis J.

Underwood, Cristina L.
Balancing Consumer Interests in a Digital Age: A New Approach to Regulating the Unauthorized Practice of Law 79:437

Vance, Holly
Protestors Have Fourth Amendment Rights, Too: In Graves v. City of Coeur d'Alene, the Ninth Circuit Clouds Clearly Established Law Governing Searches 79:753

Winn, Jane K.
Crafting a License To Know from a Privilege To Access 79:285

Winn, Peter A.
Online Court Records: Balancing Judicial Accountability and Privacy in an Age of Electronic Information 79:307

Wolcher, Louis E.
The End of Technology: A Polemic 79:331

Zorza, Richard
Some Reflections on Long-Term Lessons and Implications of the Access to Justice Technology Bill of Rights Process 79:389
Title Index

<table>
<thead>
<tr>
<th>Article</th>
<th>Authors/Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balancing Consumer Interests in a Digital Age: A New Approach to Regulating the Unauthorized Practice of Law</td>
<td>Underwood, Cristina L. 79:437</td>
</tr>
<tr>
<td>Capital Punishment, Proportionality Review, and Claims of Fairness (With Lessons from Washington State)</td>
<td>Kaufman-Osborn, Timothy V. 79:775</td>
</tr>
<tr>
<td>Conceptualizing the Right of Access to Technology</td>
<td>Horwitz, Morton J. 79:105</td>
</tr>
<tr>
<td>Crafting a License To Know from a Privilege To Access</td>
<td>Winn, Jane K. 79:285</td>
</tr>
<tr>
<td>Does Delaware’s Section 102(b)(7) Protect Reckless Directors from Personal Liability? Only if Delaware Courts Act in Good Faith</td>
<td>Berry, Matthew R. 79:1125</td>
</tr>
<tr>
<td>Government-to-Citizen Online Dispute Resolution: A Preliminary Inquiry</td>
<td>Ramasastry, Anita 79:159</td>
</tr>
<tr>
<td>Internet and the Justice System</td>
<td>Cerf, Vinton G. 79:25</td>
</tr>
<tr>
<td>Keeping Attorneys from Trashing Identities: Malpractice as Backstop Protection for Clients Under the United States Judicial Conference’s Policy on Electronic Court Records</td>
<td>Caughey, Michael 79:407</td>
</tr>
<tr>
<td>Lucky for Life: A More Realistic and Reasonable Estate Tax Valuation for Nontransferable Lottery Winnings</td>
<td>Grogan, Kyla C.E. 79:1153</td>
</tr>
<tr>
<td>Online Court Records: Balancing Judicial Accountability and Privacy in an Age of Electronic Information</td>
<td>Winn, Peter A. 79:307</td>
</tr>
<tr>
<td>Privacy as Contextual Integrity</td>
<td>Nissenbaum, Helen 79:119</td>
</tr>
<tr>
<td>Protestors Have Fourth Amendment Rights, Too: In Graves v. City of Coeur d’Alene, the Ninth Circuit Clouds Clearly Established Law Governing Searches</td>
<td>Vance, Holly 79:753</td>
</tr>
<tr>
<td>Title</td>
<td>Author(s)</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Putting Flesh on the Bones of United States v. Winans: Private Party Liability Under Treaties that Reserve Actual Fish for the Tribal Taking</td>
<td>Halm, Lindsay</td>
</tr>
<tr>
<td>Rise of the Machines: Justice Information Systems and the Question of Public Access to Court Records over the Internet</td>
<td>Silverman, Gregory M.</td>
</tr>
<tr>
<td>Settling Significant Cases</td>
<td>Seul, Jeffrey R.</td>
</tr>
<tr>
<td>Sex and the Workplace: “Consenting” Adolescents and a Conflict of Laws</td>
<td>Drobac, Jennifer Ann</td>
</tr>
<tr>
<td>Technology and the Washington State Administrative Process—Some Preliminary Notes</td>
<td>Andersen, William R.</td>
</tr>
<tr>
<td>Technology, Values, and the Justice System: Introduction</td>
<td>Alexander, Gerry</td>
</tr>
<tr>
<td>The Common Law Process: A New Look at an Ancient Value Delivery System</td>
<td>Sweeney, Dennis J</td>
</tr>
<tr>
<td>The End of Technology: A Polemic</td>
<td>Wolcher, Louis E.</td>
</tr>
<tr>
<td>Towards a Theory of Legitimate Access: Morally Legitimate Authority and the Right of Citizens to Access the Civil Justice System</td>
<td>Himma, Kenneth Einar</td>
</tr>
<tr>
<td>When Prisoners Are Weary and Their Religious Exercise Burdened, RLUIPA Provides Some Rest for Their Souls</td>
<td>Chiu, Anne Y.</td>
</tr>
<tr>
<td>Whistling in the Dark? Corporate Fraud, Whistleblowers, and the Implications of the Sarbanes–Oxley Act for Employment Law</td>
<td>Cherry, Miriam A.</td>
</tr>
<tr>
<td>Wild Dreamers: Admissibility of Dream Talk</td>
<td>Harmon, Louise</td>
</tr>
</tbody>
</table>
# Subject Index

## SUBJECT INDEX

### Accounting


### Administrative Law


### Agency

Does Delaware’s Section 102(b)(7) Protect Reckless Directors from Personal Liability? Only if Delaware Courts Act in Good Faith, Berry, Matthew R. 79:1125

### Banking and Finance


### Civil Rights


Protestors Have Fourth Amendment Rights, Too: In Graves v. City of Coeur d’Alene, the Ninth Circuit Clouds Clearly Established Law Governing Searches, Vance, Holly 79:753

When Prisoners Are Weary and Their Religious Exercise Burdened, RLUIPA Provides Some Rest for Their Souls, Chiu, Anne Y. 79:999

### Computer Law

Symposium: Technology, Values, and the Justice System. Articles by Gerry Alexander; William R. Andersen; Vinton G. Cerf; Kenneth Einar Himma; Donald J Horowitz; Morton J. Horwitz; Helen Nissenbaum; Anita Ramasasstry; Gregory M. Silverman; T.W. Small, Robert Boiko & Richard Zorza; Dennis J. Sweeney; Jane K. Winn; Peter A. Winn; Louis E. Wolcher; Richard Zorza. Comments by Michael Caughey; Christina L. Underwood 79:1

### Conflict of Laws


### Constitutional Law

Capital Punishment, Proportionality Review, and Claims of Fairness (With Lessons from Washington State), Kaufman-Osborn, Timothy V. 79:775

Protestors Have Fourth Amendment Rights, Too: In Graves v. City of Coeur d’Alene, the Ninth Circuit Clouds Clearly Established Law Governing Searches, Vance, Holly 79:753

When Prisoners Are Weary and Their Religious Exercise Burdened, RLUIPA Provides Some Rest for Their Souls, Chiu, Anne Y. 79:999

Corporations
Does Delaware’s Section 102(b)(7) Protect Reckless Directors from Personal Liability? Only if Delaware Courts Act in Good Faith, Berry, Matthew R. 79:1125


Courts


Symposium: Technology, Values, and the Justice System. Articles by Gerry Alexander; William R. Andersen; Vinton G. Cerf; Kenneth Einar Himma; Donald J Horowitz; Morton J. Horwitz; Helen Nissenbaum; Anita Ramasastry; Gregory M. Silverman; T.W. Small, Robert Boiko & Richard Zorza; Dennis J. Sweeney; Jane K. Winn; Peter A. Winn; Louis E. Wolcher; Richard Zorza. Comments by Michael Caughey; Christina L. Underwood 79:1

Criminal Law and Procedure
Capital Punishment, Proportionality Review, and Claims of Fairness (With Lessons from Washington State), Kaufman-Osborn, Timothy V. 79:775


Protestors Have Fourth Amendment Rights, Too: In Graves v. City of Coeur d’Alene, the Ninth Circuit Clouds Clearly Established Law Governing Searches, Vance, Holly 79:753

Dispute Resolution
Settling Significant Cases, Seul, Jeffrey R. 79:881

Domestic Relations
Subject Index

Employment Practice
Sex and the Workplace: “Consenting” Adolescents and a Conflict of Laws, Drobac, Jennifer Ann 79:471


Energy and Utilities Law
Putting Flesh on the Bones of United States v. Winans: Private Party Liability Under Treaties that Reserve Actual Fish for the Tribal Taking, Halm, Lindsay 79:1181

Environmental Law
Putting Flesh on the Bones of United States v. Winans: Private Party Liability Under Treaties that Reserve Actual Fish for the Tribal Taking, Halm, Lindsay 79:1181

Evidence
Wild Dreamers: Meditations on the Admissibility of Dream Talk, Harmon, Louise 79:575

Indian Law
Putting Flesh on the Bones of United States v. Winans: Private Party Liability Under Treaties that Reserve Actual Fish for the Tribal Taking, Halm, Lindsay 79:1181

Insurance Law

Jurisdiction


Jurisprudence
Settling Significant Cases, Seul, Jeffrey R. 79:881

Juveniles
Sex and the Workplace: “Consenting” Adolescents and a Conflict of Laws, Drobac, Jennifer Ann 79:471

Law and Society
Symposium: Technology, Values, and the Justice System. Articles by Gerry Alexander; William R. Andersen; Vinton G. Cerf; Kenneth Einar Himma; Donald J Horowitz; Morton J. Horwitz; Helen Nissenbaum; Anita Ramasastry; Gregory M. Silverman; T.W. Small, Robert Boiko & Richard Zorza; Dennis J. Sweeney; Jane K. Winn; Peter A. Winn; Louis E. Wolcher; Richard Zorza. Comments by Michael Caughey; Christina L. Underwood 79:1

Law Enforcement and Corrections

Protestors Have Fourth Amendment Rights, Too: In Graves v. City of Coeur d’Alene, the Ninth Circuit Clouds Clearly Established Law Governing Searches, Vance, Holly 79:753

1243
Legal Analysis and Writing
Settling Significant Cases, Seul, Jeffrey R. 79:881

Symposium: Technology, Values, and the Justice System. Articles by Gerry Alexander; William R. Andersen; Vinton G. Cerf; Kenneth Einar Himma; Donald J Horowitz; Morton J. Horwitz; Helen Nissenbaum; Anita Ramasastry; Gregory M. Silverman; T.W. Small, Robert Boiko & Richard Zorza; Dennis J. Sweeney; Jane K. Winn; Peter A. Winn; Louis E. Wolcher; Richard Zorza. Comments by Michael Caughey; Christina L. Underwood 79:1

Legislation

Motor Vehicles

Natural Resources Law
Putting Flesh on the Bones of United States v. Winans: Private Party Liability Under Treaties that Reserve Actual Fish for the Tribal Taking, Halm, Lindsay 79:1181

Organizations
Does Delaware’s Section 102(b)(7) Protect Reckless Directors from Personal Liability? Only if Delaware Courts Act in Good Faith, Berry, Matthew R. 79:1125

Practice and Procedure
Symposium: Technology, Values, and the Justice System. Articles by Gerry Alexander; William R. Andersen; Vinton G. Cerf; Kenneth Einar Himma; Donald J Horowitz; Morton J. Horwitz; Helen Nissenbaum; Anita Ramasastry; Gregory M. Silverman; T.W. Small, Robert Boiko & Richard Zorza; Dennis J. Sweeney; Jane K. Winn; Peter A. Winn; Louis E. Wolcher; Richard Zorza. Comments by Michael Caughey; Christina L. Underwood 79:1

Psychology and Psychiatry
Wild Dreamers: Meditations on the Admissibility of Dream Talk, Harmon, Louise 79:575

Religion
When Prisoners Are Weary and Their Religious Exercise Burdened, RLUIPA Provides Some Rest for Their Souls, Chiu, Anne Y. 79:999

Science and Technology
Symposium: Technology, Values, and the Justice System. Articles by Gerry Alexander; William R. Andersen; Vinton G. Cerf; Kenneth Einar Himma; Donald J Horowitz; Morton J. Horwitz; Helen Nissenbaum; Anita Ramasastry; Gregory M. Silverman; T.W. Small, Robert Boiko & Richard Zorza; Dennis J. Sweeney; Jane K. Winn; Peter A. Winn; Louis E. Wolcher; Richard Zorza. Comments by Michael Caughey; Christina L. Underwood 79:1

Securities Law
Subject Index

Sexuality and the Law
Sex and the Workplace: “Consenting” Adolescents and a Conflict of Laws, Drobac, Jennifer Ann 79:471
Wild Dreamers: Meditations on the Admissibility of Dream Talk, Harmon, Louise 79:575

State and Local Government Law

Taxation – Federal Income
Lucky for Life: A More Realistic and Reasonable Estate Tax Valuation for Nontransferable Lottery Winnings, Grogan, Kyla C.E. 79:1153

Transportation Law

Washington Law
Capital Punishment, Proportionality Review, and Claims of Fairness (With Lessons from Washington State), Kaufman-Osborn, Timothy V. 79:775
Symposium: Technology, Values, and the Justice System. Articles by Gerry Alexander; William R. Andersen; Vinton G. Cerf; Kenneth Einar Himma; Donald J Horowitz; Morton J. Horwitz; Helen Nissenbaum; Anita Ramasastry; Gregory M. Silverman; T.W. Small, Robert Boiko & Richard Zorza; Dennis J. Sweeney; Jane K. Winn; Peter A. Winn; Louis E. Wolcher; Richard Zorza. Comments by Michael Caughey; Christina L. Underwood 79:1

Women