Northern Kentucky Law Review is published four times during the academic year by students of Salmon P. Chase College of Law, Northern Kentucky University, Highland Heights, Kentucky 41099. Telephone: 859/572-5444. Facsimile: 859/572-6159. Member, National Conference of Law Reviews. All sections which appear in this issue are indexed by the H.W. Wilson Company in its Index to Legal Periodicals. Northern Kentucky Law Review articles are also available on microfilm and microfiche through University Microfilms International, Ann Arbor, Michigan.

Subscription rates are $35.00 per volume, $10.00 per individual issue. If a subscriber wishes to discontinue a subscription at its expiration, notice to that effect should be sent to the Northern Kentucky Law Review office. Otherwise, it will be assumed that renewal of the subscription is desired.

Please send all manuscripts to the address given above. No manuscript will be returned unless return is specifically requested by the author.
NORTHERN KENTUCKY LAW REVIEW

Volume 36  General Law Issue  Number 1

Editor-in-Chief
KATE MORGAN

Executive Editor
MISTY NALL

Research Editors
TROY DANIELS
DANIELLE LORENZ

Lead Articles Editor
SARAH MITCHELL

Student Articles Editor
DAWN DANLEY-NICHOLS

Spring Symposium Editor
BRYCE RHoades

Fall Symposium Editor
RHONDA SCHECHTER

Kentucky Survey Issue Editor
AMANDA McFARLAND

General Symposia Editor
CHARLIE VANCE

Special Topics Editor
AARON STACK

Associate Editors
ROB ALTMAN
MIKE ANGLIN
KRIS BRANDENBURG
CHAD ECKHARDT
CHRISTINE ELECH
ASHLEY ENGLAND-HUFF
CHRISTIAN FANNIN
BRAD GIBSON
MEGAN HALL
MICHELLE HOFF
KERRY HOLLERAN
JOANNA HORTILLOSA
DANIEL HUNT
BEN LEWIS
EMILY LITZINGER
BRIAN O’CONNOR

Faculty Advisor
PROFESSOR DAVID ELDER

Technical Advisor
PROFESSOR RICHARD BALES
SALMON P. CHASE COLLEGE OF LAW

DENNIS HONABACH, A.B., J.D.
Dean & Professor of Law

LAWRENCE ROSENTHAL, B.A., J.D., LL.M.
Associate Dean for Academics

JENNY YOUNG
Registrar of the College of Law

LISA MOORE, B.A., J.D.
Director of Career Services

MOLLY TAM
Associate Dean for Student Affairs and
Admissions & Director of Pro Bono Services

ASHLEY GRAY, B.A., J.D.
Director of Admissions

KAREN OGBURN, B.A., M.P.A.
Associate Dean for Administration

SHERRIE TURNER, B.S.
Budget Analyst

DAVID H. MACKNIGHT, B.S., J.D.
Associate Dean for Law School Advancement

MICHAEL WHITEMAN, B.A., M.S.L.S., J.D.
Associate Dean for Library Services and
Information Technology & Professor of Law

FACULTY

RICHARD A. BALES, B.A., J.D.
Professor of Law & Associate Dean of Faculty Development

IAIN BARKSDALE, B.A., M.PHIL., M.L.S., J.D.
Assistant Professor of Library Services
Assistant Director of Information Technology

DONNA S. BENNETT, A.B., M.A., M.S.L.S.
Professor of Library Services &
Assistant Director for User Services

JOHN M. BICKERS, A.B., J.D., LL.M.
Assistant Professor of Law

ROGER D. BILLINGS, JR., A.B., J.D.
Professor of Law

CAROL BREDEMEYER, A.B., M.S.L.S.
Professor of Library Services &
Assistant Director of Faculty Services

ALEA BROWN, B.A., J.D.
Visiting Professor of Legal Writing

DAVID A. ELDER, A.B., J.D., LL.M.
Professor of Law

NANCY L. FIKR B.S., J.D., LL.M.
Professor of Law

CAROL FURNISH, B.A., M.L.S.
Professor of Library Services &
Assistant Director for Instruction & Outreach

CHRISTOPHER GUINELLO, B.A., J.D., LL.M.
Associate Professor of Law

THOMAS E. HEARD, B.A., M.L.S., J.D.
Associate Professor of Library Services &
Associate Director for Library Services

DAVIDA H. ISAACS, B.A., J.D.
Associate Professor of Law

EMILY JANOSKI-HAEHLIN, B.S., J.D., M.L.S.
Assistant Professor of Library Services
Assistant Director for Research and Online Services

KATHLEEN JOHNSON, B.A., J.D.
Assistant Professor &
Associate Director of the Chase Center for Excellence in Advocacy

JENNIFER JOLLY-RYAN, B.A., J.D.
Professor of Legal Writing

KENNETH D. KATKIN, A.B., J.D.
Professor of Law

DONALD K. KAZEE, B.A., M.A., J.D.
Associate Professor of Legal Writing

JENNIFER KREIDER, B.A., J.D.
Associate Professor of Law

SHARLENE LASSITER-BOLTZ, B.A., J.D.
Professor of Law

ALBERTO LOPEZ, B.S., M.S., J.D., J.S.M.
Professor of Law

MICHAEL J.Z. MANNHEIMER, B.A., J.D.
Associate Professor of Law

BARBARA MACFARLAND
Acting Director of Student Success Initiatives
Assistant Professor

LIUBOMIR NACEV, B.A., J.D., LL.M.
Professor of Law

ADRIENNE NOBLE NACEV, B.A., J.D.
Visiting Assistant Professor of Legal Writing

SHERRY PORTER,
Director for the Center for Excellence in Transactional Law

LOWELL F. SCHECHTER, B.A., J.D.
Professor of Law

DAVID SINGLETON, B.A., J.D.
Visiting Professor of Law

WILLIAM SJOSTROM, B.A., J.D.
Professor of Law

PHILLIP SPARKES, B.S., J.D., LL.M.
Assistant Professor of Law &
Director of Chase Local Government Law Center

MARK M. STAVSKY, B.A., J.D., LL.M.
Professor of Law

HENRY L. STEPHENS, JR., B.A., J.D.
Executive Director of Chase Center for Excellence in Advocacy

JOHN T. VAALUARI, B.A., J.D.
Professor of Law
CARYL A. YZENBAARD, A.B., J.D.  
Professor of Law

CLAUDIA ZAHER, B.A., M.S.L.S.  
Associate Professor of Library Services  
Assistant Director for Collection Development  
Acquisitions and Serials

EMERITI FACULTY
CAROL B. ALLRED, B.A., M.S.L.S., J.D.
ROBERT M. BRATTON, A.B., M.A., J.D.
EDWARD P. GOGGIN, J.D., LL.M.
W. JACK GROSSE, B.S.C., M.B.A., J.D.
MARTIN J. HUELSMAN, B.B.A., J.D.
WILLIAM R. JONES, B.S., M.A., J.D.
KAMILA M. MAZENAC, A.B., J.D., LL.M.
FREDERICK R. SCHNEIDER, B.A., J.D.
ROBERT L. SEAVER, A.B., J.D.
DAVID C. SHORT, B.A., J.D., LL.M.

ADJUNCT FACULTY
HON. GREGORY M. BARTLETT, A.B., M.A., J.D.
HON. WILLIAM O. BERTELSMAN, B.S., J.D.
BRYANT BREWER, B.A., J.D.
LOUISE BROCK
SHAWN P. BURTON
KELLY CANTRALL, B.A., J.D.
LISA DAVIDSON, B.A., J.D.
KRISTI DAVIDSON, B.A., J.D.
REBECCA DILORETO
DOREENA FOX, B.S., J.D.
ROBERT A. GOERING, B.A., J.D.
HON. DANIEL T. GUIDUGLI, B.A., J.D.
MICHAEL HAWKINS, B.A., J.D.
HON. STEVEN R. JAEGE
ARTICLES

De-Clothing Sex-Based Classifications — Same-Sex Marriage is Just the Beginning: Achieving Formal Sex Equality in the Modern Era
   Catherine Jean Archibald .............................................................................................................. 1

Coupling Responsibility with Liability: Why Institutional Review Board Liability is Good Public Policy
   Danielle C. Beasley .......................................................................................................................... 45

The Doctrine of Other Wrongs: A Framework for Punishing Civil Recidivism
   Miriam R. Coles .............................................................................................................................. 67

Frankfurter’s Gloss Theory, Separation of Powers, and Foreign Investment
   Catherine H. Gibson ....................................................................................................................... 103

NOTES

Free Enterprise Fund v. Public Company Accounting Oversight Board: The Decision That Corporate America May Forever Be Waiting For
   Chad N. Eckhardt ............................................................................................................................ 143

A Cosmopolitan Approach to Treaty Interpretation: Why Service By Postal Channels Should Be Permitted Under the Hague Convention
   Christine A. Elech .......................................................................................................................... 163

Doninger v. Niehoff: “Tinker is Online and in Trouble”
   Bradley M. Gibson .......................................................................................................................... 185
ERRATA

After Issue 4 of Volume 35 was published, it was brought to our attention that the Northern Kentucky Law Review mistakenly reversed the order of the authors for the article, Does Stigmatizing Sex Offenders Drive Down Reporting Rates? Perverse Effects and Unintended Consequences. The lead author of the aforementioned article is Heather R. Hlavka, and the co-author is Chris Uggen.

Our apologies,

Katie R. Morgan
Editor-in-Chief
Northern Kentucky Law Review
2008-2009

DE-CLOTHING SEX-BASED CLASSIFICATIONS - SAME-SEX MARRIAGE IS JUST THE BEGINNING: ACHIEVING FORMAL SEX EQUALITY IN THE MODERN ERA

Catherine Jean Archibald

Table of Contents

ABSTRACT .............................................................................................................2
INTRODUCTION ......................................................................................................3
I. SEX DISCRIMINATION SHOULD BE JUDGED WITH STRICT SCRUTINY .... 5
   A. Sex and race discrimination are judged by different levels of scrutiny, which results in different levels of protection ........ 5
   B. Sex and race are similar traits and both should be protected by strict scrutiny ................................................................. 9
II. SEX-BASED CLASSIFICATIONS CAUSE NEGATIVE EFFECTS .............13
   A. There is inherent indignity and injustice in being classified by sex ......................................................................................... 14
   B. Sex-based classifications reinforce sex-based stereotypes......17
      1. The same-sex marriage restriction reinforces sex stereotypes ............................................................................................. 18
      2. Sex segregation in sports reinforces sex stereotypes ........ 18
      3. Sex segregation in bathrooms reinforces sex stereotypes ........................................................................................................ 19
      4. Sex-specific pronouns reinforce sex stereotypes .................. 20
   C. Sex-based classifications exacerbate discrimination against and otherwise cause harm to people who do not conform to sex stereotypes ................................................................................................. 21
      1. Outlawing same-sex marriage harms LGBT people .......... 24
      2. Sex-segregated sports harm “sissy” boys and LGBT people .................................................................................................... 25

* Catherine Jean Archibald is a graduate of the joint-degree J.D.-L.L.B. program run by Michigan State University College of Law and the University of Ottawa Faculty of Law, and has a bachelor’s degree in Ecology and Evolutionary Biology from Princeton University. Currently, Catherine works as a New York based associate at the international law firm of Dewey & LeBoeuf LLP. The views and opinions expressed in this article are solely those of the author and do not necessarily represent the views of Dewey & LeBoeuf LLP. Catherine can be reached at catherine.archibald@gmail.com. The author would like to thank Melanie Jacobs and the Michigan State Law Review for their help in writing the first draft of this paper. Special thanks also to N. Kathleen (“Sam”) Banks for all the comments, suggestions, and support.
3. Sex-segregated bathrooms harm those who fail to “look their sex” ................................................................. 25
4. Sex-specific pronouns harm transgender people .......... 25

III. If Courts Ruled That Sex-Based Classifications Must Be Judged with Strict Scrutiny, Several Societal Changes Would Occur ................................................................. 26
   A. Same-sex marriage would be permitted .................. 26
      1. “Separate but equal” treatment is not equal ........... 27
      2. Purported state interests in procreation do not meet strict scrutiny .......................................................... 30
      3. Purported state interests in the best interest of the child do not meet strict scrutiny .................................. 31
      4. Purported state interests in religion do not meet strict scrutiny ................................................................. 31
   B. Sports in schools would be integrated ....................... 33
      1. Safety concerns do not meet strict scrutiny ............. 35
      2. Concerns that women will get squeezed out of sports do not meet strict scrutiny .................................. 36
      3. Redressing past discrimination against women does not meet strict scrutiny ........................................ 37
   C. Bathrooms would be integrated ................................. 38
      1. Comfort, convenience, and privacy concerns do not meet strict scrutiny .................................................. 38
      2. Safety concerns do not meet strict scrutiny ............. 39
   D. Language would change by the use of sex-neutral pronouns .... 40

IV. Under the Proposed Regime, Few Sex Classifications Would Remain .......................................................... 41
   A. De facto segregation may still exist .......................... 41
   B. Rare situations may justify sex-based classifications .......... 42

CONCLUSION .................................................................................................................. 42

ABSTRACT

Last summer the highest court of California extended marriage-rights to same-sex couples, becoming the second U.S. state to do so, after Massachusetts. Shortly afterwards, New York started recognizing same-sex marriages performed outside the state and Connecticut’s highest court recognized the fundamental right of same-sex couples to get married. A few months later, Iowa’s Supreme Court declared its same-sex marriage ban unconstitutional, and Vermont became the first state legislature to legalize same-sex marriage. With these events, within just over five years time, the percentage of Americans living in states that
recognize same-sex marriage went from 0% to over 20%. As individual judges correctly noted in three of the six high court marriage decisions of these states, recognizing same-sex marriage rights is an important step towards achieving formal sex equality. After all, men and women should have equal rights and opportunities. Therefore, if a man has the right to marry a woman, then a woman must also have the right to marry a man, and vice versa.

This Article argues that not only same-sex marriage, but also co-ed sports teams, co-ed bathrooms, and sex-neutral pronouns are current constitutional imperatives under the Equal Protection Clause of the Fourteenth Amendment, necessary to achieving formal sex equality, because any sex-based classification should be judged with strict scrutiny. Sex-based classifications, like race-based classifications, are based on an immutable physical trait with little relevance to most sex-segregated rights. Furthermore, sex-based classifications cause negative effects in society, such as increased sex stereotyping and discrimination against those who do not conform to sex-stereotypes. These current problems would be ameliorated by judging sex-based classifications with strict scrutiny and eliminating any remaining sex-based differences in constitutional rights.

INTRODUCTION

Last summer the highest court of California, the most populous U.S. state,\(^1\) extended marriage-rights to same-sex couples\(^2\). It was the second U.S. state to do so, preceded only by Massachusetts.\(^3\) Shortly afterwards, New York started recognizing same-sex marriages performed outside the state,\(^4\) and Massachusetts began allowing non-resident same-sex couples to get married within its borders.\(^5\) A few months after that, the highest courts in Connecticut and Iowa recognized the fundamental right of same-sex couples to marry, and Vermont became the

---


first state legislature to legalize same-sex marriage.\textsuperscript{6} Thus, in just slightly over five years time, the percentage of Americans living in states that recognize same-sex marriage went from 0% to over 20%.\textsuperscript{7}

As individual judges correctly noted in three of the six high court same-sex marriage decisions of these states, recognizing same-sex marriage rights is an important step towards achieving formal sex equality\textsuperscript{8} under the Equal Protection Clause.\textsuperscript{9} Under the Equal Protection Clause, men and women should have equal


\textsuperscript{8} Formal equality involves treating individual men and women equally, without meting out a difference in rights based on sex. Formal equality is the focus of this Article and is an important starting point to ensure that people are treated fairly under the law. See United States v. Vue, 13 F.3d 1206, 1213 (8th Cir. 1994) (“Formal equality before the law is the bedrock of our legal system . . . ”). Often, additional measures are needed to ensure “substantive equality,” or equality of opportunity for people who come from different starting points. Substantive equality is relevant and discussed in this Article in the context of sports. See infra Part III:B:2.

\textsuperscript{9} See Hernandez v. Robles, 855 N.E.2d 1, 29-30 (N.Y. 2006) (Kaye, C.J., dissenting); Goodridge v. Dep’t of Public Health, 798 N.E.2d 941, 972-73 (Mass. 2003) (Greaney, J., concurring); Baker v. State, 744 A.2d 864, 905 (Vt. 1999) (Johnson, J., concurring in part and dissenting in part) (recognizing that the marriage statute classifies on the basis of sex). Numerous judges on other courts have come to the same conclusion, although this fact has only been recognized by a high court majority on one occasion. See, e.g., Baehr v. Lewin, 852 P.2d 44, 67 (Haw. 1993) (holding that the opposite-sex marriage requirement classifies based on sex and implicates the equal protection clause); Andersen v. King County, 138 P.3d 963, 1037 (Wash. 2006) (en banc) (Bridge, J., concurring in dissent) (explaining that the opposite-sex marriage requirement discriminates based on sex because “the only thing preventing plaintiff Heather Andersen from marrying her partner, Leslie Christian, is the fact that Andersen is a woman. Andersen should no more readily be prohibited from marrying her partner than she is from voting for president or practicing law.”); Varnum v. Brien, No. CV5965, 2007 WL 2468667 (Iowa Dist.) (Trial Order) (finding that the opposite-sex marriage requirement impermissibly discriminates based on sex because “[e]ach plaintiff would have been able to marry his or her partner had the Plaintiff been of a different sex”). See also Jennifer Levi, Toward a More Perfect Union: The Road to Marriage Equality for Same-Sex Couples, 13 WIDENER L.J. 831, 843 (2004) (allowing opposite-sex marriage but not same-sex marriage is sex discrimination because if a person in a gay male relationship would be allowed to marry his partner if he were female, then the exclusion is based upon his sex); Andrew Koppelman, Why Discrimination Against Lesbians and Gay Men Is Sex Discrimination, 69 N.Y.U. L. REV. 197, 219 (1994) (arguing that the idea that the behavior of sex with women is appropriate for one sex and not the other, and vice versa, is based on traditional sex-based gender roles and is therefore sex-based discrimination); Stephen Clark, Same-Sex but Equal: Reformulating the Miscegenation Analogy, 34 Rutgers L.J. 107, 115-16 (2002); Pamela S. Katz, The Case for Legal Recognition of Same-Sex Marriage, 8 J.L. & Pol’y 61, 88-92 (1999); Jeffrey Hubins, Proposition 22: Veiled Discrimination or Sound Constitutional Law?, 23 WHITTIER L. REV. 239, 258-60 (2001). But see infra notes 173-76 and accompanying text; Lynn D.
rights and opportunities. Therefore, if a man has the right to marry a woman, then a woman must also have the right to marry a woman, and vice versa.

Thus, legalizing same-sex marriage advances the goal of formal sex equality and equal protection under the law because allowing same-sex marriage gives people rights and duties based on their personhood, not on their sex. But even with the increase in legalization of same-sex marriage in the United States, there is still much progress to be made in marriage rights and in other areas before formal sex equality is achieved.

The Article focuses on what the author believes are some of the most prevalent examples of legally-sanctioned, harm-causing, sex-based classifications\(^\text{10}\) in American society today. These are sex segregation in: 1) marriage rights; 2) school sports; 3) public bathrooms; and 4) the use of sex-specific pronouns.\(^\text{11}\) Part I describes current United States Supreme Court jurisprudence on sex-based classifications and why sex discrimination should be judged by strict scrutiny, the same level of scrutiny used for race discrimination. Part II describes the harm sex classifications cause society. Part III describes the consequences that would result if sex classifications were judged with strict scrutiny. Finally, Part IV describes when sex classifications would be permitted even under a strict scrutiny regime.

I. SEX DISCRIMINATION SHOULD BE JUDGED WITH STRICT SCRUTINY

Currently, sex discrimination is judged more leniently than race discrimination. This should change, as both types of discrimination are similar and deserve the same level of scrutiny.

A. Sex and race discrimination are judged by different levels of scrutiny, which results in different levels of protection

Joe and Mary are in love and want to get married. They apply for a marriage license but are denied because Joe is black and Mary is white.\(^\text{12}\) Joe and Tom are best friends and like spending time together. Unfortunately, they must play on

---

\(^{10}\) In this Article, the terms “sex classification,” “sex discrimination,” and “sex segregation” are used interchangeably in that in each instance a person is provided different rights under the law, based on sex alone. See BLACK’S LAW DICTIONARY 500 (8th ed. 2004) (defining discrimination as “differential treatment”).

\(^{11}\) The author acknowledges that there are other examples of sex-based classifications not explicitly addressed in this Article, such as sex segregation in changing rooms and the military. While not addressed, a similar analysis and criticism would be given to these other forms of sex segregation.

\(^{12}\) See Loving v. Virginia, 388 U.S. 1, 12 (1967) (finding laws forbidding interracial marriage invalid).
separate baseball teams, sit in separate parts of their favorite restaurant, drink from separate water fountains, and go to separate bathrooms, all because Joe is black and Tom is white.

These restrictions based on race were once enforced by law and supported by a society with a history of race discrimination and stereotypes; however, they no longer exist due to the United States Supreme Court’s rulings under the Equal Protection Clause of the Fourteenth Amendment. The Equal Protection Clause provides that “No State shall . . . deny to any person within its jurisdiction the equal protection of the laws.” Accordingly, the law no longer permits these types of race segregation or discrimination. By contrast, when it comes to sex, a different rule applies.

Sam and Julie are in love and want to get married. They apply for a marriage license but are denied because they are both women. Sam and Adam are best friends and like spending time together. Unfortunately they must play on separate baseball teams and go to separate bathrooms, because Sam is female and Adam is male. These restrictions based on sex are still enforced today by law, supported by a society with a history of sex discrimination and stereotypes.

---

15. See Blacks United for Lasting Leadership, Inc. v. City of Shreveport, 571 F.2d 248, 257 (5th Cir. 1978) (Wisdom, J., dissenting) (describing a city’s segregation system, where blacks had to use separate water fountains from whites).
17. See supra notes 12-16.
18. See, e.g., supra notes 12-16; Watson v. City of Memphis, 373 U.S. 526, 539 (1963) (outlawing racial segregation in public parks and recreational facilities); Brown v. Bd. of Educ., 347 U.S. 483, 495 (1954) (outlawing racial segregation in public schools); City of St. Petersburg v. Alsup, 238 F.2d 830, 832 (5th Cir. 1956) (outlawing racial segregation at public beaches and swimming pools). See also Garner v. Louisiana, 368 U.S. 157, 184 (1961) (Douglas, J., concurring) (concluding that because of the Equal Protection Clause, race is an “impermissible classification when it comes to parks or other municipal facilities”).
19. U.S. CONST. amend. XIV, § 1. Many state constitutions echo this equal protection requirement. See, e.g., N.Y. CONST. art. I, § 11 (“No person shall be denied the equal protection of the laws of this state or any subdivision thereof.”); HAW. CONST. art. I, § 5 (“No person shall be . . . denied the equal protection of the laws . . . .”). While the Fourteenth Amendment on its face only applies to states, not the federal government, the Supreme Court has held that the Due Process Clause of the Fifth Amendment, which applies to the federal government, includes the duties of the Equal Protection Clause of the Fourteenth Amendment. See Weinberger v. Wiesenfeld, 420 U.S. 636, 638 n.2 (1975) (noting that the approach to Fifth Amendment equal protection claims is the same as Fourteenth Amendment claims); Bolling v. Sharpe, 347 U.S. 497, 499 (1954) (unjustifiable discrimination violates the federal Due Process Clause).
20. See infra Parts II.B.1, III.A.
21. See infra Parts II.B.2, III.B.
22. See infra Parts II.B.3, III.C.
This difference in treatment for race and sex is due to the different treatment given to each by the United States Supreme Court. The Court has held that government race-based classifications must pass a “strict scrutiny” test to be upheld; whereas sex-based classifications need only pass the less stringent “intermediate scrutiny” test. In practical terms, this means that a white male receives more protection from race discrimination than a female receives from sex discrimination. The “strict scrutiny” test used for racial classifications requires that the classification be narrowly tailored to meet a compelling state interest. By contrast, the “intermediate scrutiny” test, used for sex classifications, uses weaker language, requiring only that the classification be substantially related to an important government interest.

Under the “strict scrutiny” test, a “compelling” governmental interest justifying a race-based classification is found only in the most rare and unusual of circumstances. It is found only in the need to protect human life, in narrow circumstances to diversify an academic institution, and in providing remedies for specific instances of past discrimination.

By contrast, under the “intermediate scrutiny” test, what qualifies as an “important” government interest “substantially” related to a sex-based classification is ambiguous and encourages inconsistent judicial interpretation.

23. Hereinafter, when sex or race-based classifications are mentioned, it is implied that they are made by the government, unless otherwise stated. The legality of privately-made race and sex classifications is outside the scope of this paper.


25. See Adarand, 515 U.S. at 227 (race based distinctions made by the government, which disadvantage whites, are subject to strict scrutiny). See also Deborah L. Brake, Sex as a Suspect Class: An Argument for Applying Strict Scrutiny to Gender Discrimination, 6 SETON HALL CONST. L.J. 953, 961 (1996) (“[W]hite men . . . have greater legal protection from reverse discrimination by affirmative action programs than women have from government-sponsored discrimination against women.”).


28. See City of Richmond v. J.A. Croson Co., 488 U.S. 469, 521 (1989) (Scalia, J., concurring) (stating that nothing more than a “social emergency rising to the level of imminent danger to life and limb - for example, a prison race riot, requiring temporary segregation of inmates, . . . can justify [race segregation]”).

29. Compare Grutter v. Bollinger, 539 U.S. 306, 341-43 (2003) (race can be used as a “plus” factor in admitting students to a university to increase diversity of the student body, but in twenty-five years this will not be necessary), with Parents Involved in Cmty. Sch. v. Seattle Sch. Dist. No. 1, 127 S. Ct. 2738, 2768 (2007) (plurality opinion) (race cannot be used to deny or admit students in a public school district).

30. See Adarand, 515 U.S. at 239 (Scalia, concurring) (arguing that government cannot “make up” for past discrimination by discriminating in the opposite direction; there are no “creditor” or “debtor” races; however, individuals who have been harmed by specific acts of racial discrimination should be compensated).

31. This ambiguity arises because the standard is less strict than “strict scrutiny,” which few discriminatory laws pass, and yet stricter than “rational basis” scrutiny, which most discriminatory laws pass. See Note, Making Sense of Hybrid Speech: A New Model for Commercial Speech and
For example, some courts have found that there is an important government interest that allows girls to be excluded from boys’ sports teams, while other courts have found that there is not.32

Current United States Supreme Court rulings, that sex discrimination need only be judged with intermediate scrutiny, encourage judges’ personal opinions of appropriate sex roles to inform their decisions and are inadequate to eradicate sex stereotypes in law.33 In cases such as Craig v. Boren and United States v. Virginia, laws made on the basis of stereotypes, that women are more temperate than men and that men are better able to handle rigorous physical and mental training, were struck down.34 However, in other cases, such as Michael M. v. Superior Court and Nguyen v. Immigration and Naturalization Service, laws made on the basis of stereotypes that women carry all the burdens of pregnancy and childrearing and that among unmarried couples, women will be the ones to develop a relationship with and raise any children, were upheld.35

Because sex is not afforded as much protection as race, many sex-based classifications are currently tolerated where the same classification based on race would not be tolerated. While it is inexcusable to separate people by race in order to determine who may marry whom,36 the federal government, and almost

---

Expressive Conduct, 118 Harv. L. Rev. 2836, 2854 (2005) (explaining that judges have the most discretion with intermediate scrutiny). See also Virginia, 518 U.S. at 559 (Rehnquist, C.J., concurring) (stating that ‘terms like ‘important governmental objective’ and ‘substantially related’ are hardly models of precision’); Brake, supra note 25, at 958 (stating that “the intermediate scrutiny standard provides insufficient guidance and leaves broad discretion with individual judges”); Holning Lau, Sexual Orientation & Gender Identity: American Law in Light of East Asian Developments, 31 Harv. J. L. & Gender 88-89 (2008) (criticizing the three-tiered approach used in U.S. equal protection jurisprudence). Even if strict scrutiny is not adopted, the sex-based classifications discussed in this article should be struck down, under intermediate scrutiny. Cf. Varnum v. Brien, No. 07-1499, 2009 Iowa Sup. LEXIS 31 (Iowa Apr. 3, 2009) (finding that Iowa’s same-sex marriage statute could not withstand even intermediate scrutiny, and thus finding it unnecessary to determine whether strict scrutiny should be used).


33. See infra note 35 and accompanying text. See also Ann Elizabeth Mayer, Reflections on the Proposed United States Reservations to CEDAW: Should the Constitution be an Obstacle to Human Rights? 23 Hastings Const. L.Q. 727, 785 (1996) (stating that the intermediate scrutiny standard “allows judges to factor in sex stereotypes in deciding whether discrimination is unconstitutional”).


36. See Loving v. Virginia, 388 U.S. 1, 11-12 (1967) (holding that racial classifications for purposes of marriage is unconstitutional).
all state governments, separate people by sex in order to determine just that.\textsuperscript{37} Jim Crow laws mandating separate public facilities for blacks and whites were abolished decades ago,\textsuperscript{38} but today, a majority of the public must still frequent sex-segregated bathrooms in public places.\textsuperscript{39} While separating children by race in order to play sports in school would be considered outrageous, separating children by sex happens daily in schools across the country.\textsuperscript{40} Finally, while labeling a child by race in a non-race related classroom discussion would be deemed racist, labeling children by their sex in a non-sex related discussion happens every day in schools across the country.\textsuperscript{41}

\textbf{B. Sex and race are similar traits and both should be protected by strict scrutiny}

Both race-based and sex-based classifications should be judged with the highest level of scrutiny. The Equal Protection Clause states that “No State shall . . . deny to any person within its jurisdiction the equal protection of the laws.”\textsuperscript{42}

\textsuperscript{37} See, e.g., 1 U.S.C. § 7 (2006) (“[T]he word ‘marriage’ means only a legal union between one man and one woman as husband and wife, and the word ‘spouse’ refers only to a person of the opposite sex who is a husband or a wife.”) (part of the federal Defense of Marriage Act (DOMA) that became law in 1996); Mich. Const. art. I, § 25 (2006) (“[T]he union of one man and one woman in marriage shall be the only agreement recognized as a marriage or similar union for any purpose.”); Alaska Stat. § 25.05.013(b) (2005) (“A same-sex relationship may not be recognized by the state as being entitled to the benefits of marriage.”); Ga. Code Ann. § 19-3-3.1(a) (2006) (“It is declared to be the public policy of this state to recognize the union only of man and woman. Marriages between persons of the same sex are prohibited in this state.”). As of this writing, there are only six states in the United States that recognize same-sex marriage: Massachusetts, California, New York, Connecticut, Iowa, and Vermont, and only four states that perform same-sex marriage: Massachusetts, Connecticut, Iowa, and Vermont. See \textit{In re Varnum v. Brien}, No. 07-1499, 2009 Iowa Sup. LEXIS 31 (Iowa Apr. 3, 2009); Marriage Cases, 183 P.3d 384, 399 (Cal. 2008); Kerrigan v. Comm'r of Pub. Health, 957 A.2d 407, 412 (Conn. 2008); Goodridge v. Dep't of Pub. Health, 798 N.E.2d 941, 948 (Mass. 2003); Egelko & Wildermuth, supra note 2; Kerrigan v. Comm'r of Pub. Health, supra note 4; An Act Relating to Civil Marriage, supra note 6. Internationally, five countries allow same-sex marriage: the Netherlands, Belgium, Spain, Canada, and South Africa. See Netherlands; Civil Code, Art. 30 (as amended Dec. 21, 2000); Belgium: Civ. code, art. 143 (as amended Feb. 13, 2003); Spain: C.C., art. 44 (as amended July 1, 2005); Canada: Civil Marriage Act, 2005 S.C., ch. 33 (Can.); South Africa: Civil Union Act 17 of 2006, s. 11.

\textsuperscript{38} See, e.g., Turner v. City of Memphis, 369 U.S. 350, 352-54 (1962) (ordering end to state policy of racial segregation in publicly operated facilities); Garner v. Louisiana, 368 U.S. 157, 184 (1961) (Douglas, J., concurring) (concluding that because of the Equal Protection Clause, race is an “impermissible classification when it comes to parks or other municipal facilities”).

\textsuperscript{39} See Goins v. West Group, 635 N.W.2d 717, 723 (Minn. 2001) (sex-segregated bathrooms are common).

\textsuperscript{40} See O'Connor v. Bd. of Educ., 449 U.S. 1301, 1306-08 (1980) (holding that a girl may be excluded from boys' basketball team even when the boys' team plays on a higher level than the girls' team and the girl is as good as the players on the boys' team); see also Dana Robinson, \textit{A League of Their Own: Do Women Want Sex-Segregated Sports?}, 19 J. CONTEMP. LEGAL ISSUES 321, 322 (1998) (explaining that current equal protection jurisprudence under the Fourteenth Amendment often does not help women who want to be integrated into mainstream sports).

\textsuperscript{41} Children are identified by sex daily when they are referred to by sex-specific pronouns. See infra Part III:D.

\textsuperscript{42} U.S. Const. amend. XIV, § 1.
This text provides no basis for a difference in protection between race and sex, both physical traits. Although the Equal Protection Clause was enacted in order to protect the rights of African-American males, the clause is broadly worded, and has since been used to protect against discrimination based not only on race, but also national origin, religion, sex, mental disability, illegitimacy, and sexual orientation. Thus, since the Equal Protection Clause is broadly worded and is used to protect many different personal characteristics, any difference in level of protection between two characteristics should be well justified.

However, the United States Supreme Court has never explicitly justified the difference in protection given for race and sex. One author suggests that the intermediate scrutiny standard, which was first developed to judge sex discrimination, was simply a compromise between those Justices who wanted sex discrimination judged under rational basis, and those Justices who wanted it judged under strict scrutiny.

Nonetheless, perhaps a justification for the difference in treatment for race discrimination and sex discrimination can be gleaned by examining the factors the Court has listed as relevant in determining what level of scrutiny should be used to protect a particular characteristic. These factors are: 1) whether the characteristic determines the individual’s ability to perform, or to participate in, or contribute to, society; 2) whether the characteristic is immutable; 3) whether persons with the characteristic have historically suffered from discrimination; 4) whether the group is “politically powerless.”

---

43. See McLaughlin v. Florida, 379 U.S. 184, 192 (1964); Mayer, supra note 33, at 772.
44. See Korematsu v. United States, 323 U.S. 214, 216 (1944).
50. The lowest level of scrutiny under the Equal Protection clause is the “rational basis” level of scrutiny, which simply requires that the challenged law be rationally related to a legitimate government interest. See Ry. Express Agency v. New York, 336 U.S. 106, 110 (1949).
53. See Frontiero, 411 U.S. at 686.
54. See Cleburne, 473 U.S. at 441; Murgia, 427 U.S. at 313; Frontiero, 411 U.S. at 684-85.
Looking at these four factors, sex and race are equally deserving of protection against discrimination under the first three factors. Neither sex nor race determines a person’s ability to contribute to society.\(^{56}\) Both sex and race are equally immutable.\(^{57}\) And it is indisputable that people in this country have historically suffered great discrimination because of their race and because of their sex.\(^{58}\)

Only in the fourth factor, “political powerlessness,” is there a significant difference between groups historically disadvantaged because of race and groups disadvantaged because of sex.\(^{59}\) Racial minorities may be discrete and insular minority populations, unlike the population of either men or women.\(^{60}\) However, this factor should not be used by courts in determining whether a group should be protected by strict scrutiny.

The “political powerless” factor muddles the functions of the judicial branch with the legislative branch. The judicial branch has a duty to uphold the constitution, no matter how unpopular or popular it is to do so,\(^{61}\) while the legislative branch enacts laws based on the popular vote. Thus, whether or not the discriminated class is able to mobilize politically should be irrelevant to whether that class deserves protection under the Equal Protection Clause before a court of law. Put simply, a person seeking equal protection under the law, a constitutional right, should not have the burden of mobilizing the entire country before obtaining that right. This is the essence of constitutional rights, which ensure that the tyranny of the majority (including parts of majority that are members of the claimant’s class) will be avoided. Indeed, a majority of people may be complacent about securing specific equality rights, even for themselves, as is the case for sex equality.\(^{62}\) Additionally, a particular instance of discrimination may negatively affect only a segment of the class discriminated against, making it hard to mobilize politically.\(^{63}\)

The United States Supreme Court has indirectly agreed with this conclusion in its decisions to protect members of the European-American racial group from

\(^{56}\) See *Frontiero*, 411 U.S. at 686.

\(^{57}\) Although some people opt to change their sex through surgery, this is a difficult, expensive procedure, and should not be required before equal protection is conferred. See Lynn Conway, *How Frequently Does Transsexualism Occur?*, available at http://ai.eecs.umich.edu/people/conway/TS/TSprevalence.html (last visited Dec. 17, 2008). See also supra note 102.

\(^{58}\) See *Sail'er Inn, Inc. v. Kirby*, 485 P.2d 529, 540-41 (Cal. 1971) (en banc).

\(^{59}\) Therefore, this is the only factor that shall be critiqued in this Article, even though some of the other factors may also merit critique.

\(^{60}\) See *Carolene Products Co.*, 304 U.S. at 152-53 n.4.

\(^{61}\) See *People ex rel. N.R.*, 139 P.3d 671, 681 (Colo. 2006) (en banc) (under the separation of powers doctrine, courts must be independent and enforce the constitution even if their decisions are unpopular); *Fong v. Am. Airlines, Inc.*, 431 F. Supp. 1334, 1340 (N.D. Cal. 1977) (court has the “duty to make hard and unpopular decisions where it considers them correct”).

\(^{62}\) For example, many people with whom I have spoken are not concerned about the sex classifications I describe in this Article.

\(^{63}\) See infra Part II.C. (Many sex discriminations disproportionately harm a subclass of the group classified by sex).
race discrimination using strict scrutiny.64 By its decisions, the Court impliedly acknowledges that a group need not be politically powerless before it is protected from discrimination by strict scrutiny because European-Americans are the most politically powerful group in the United States.65 European-Americans are protected by a strict scrutiny standard from racial discrimination, even though they have not suffered historic discrimination and are not a discrete and insular minority.66 Thus, because European-Americans do not need to mobilize politically every time a law is passed that discriminates against them, neither should other politically powerful groups.

The California Supreme Court has agreed with this reasoning in its recent decision on same-sex marriage. In discussing the factors relevant to determining whether a characteristic should be protected by strict scrutiny, the Court dismissed the importance of “current political powerless,” explaining that if that were required for a group to achieve suspect classification status, race and religion would no longer exist as suspect classifications.67

Sex is a characteristic that, like race, cannot definitively predict an individual’s talents, preferences, or abilities.68 Additionally, sex is a

65. U.S. Census Bureau, http://factfinder.census.gov/home/saff/main.html?_lang=en (follow “People” hyperlink; then follow “Race and Ethnicity” hyperlink; then follow “Race”) (showing white Americans as the majority racial group) (last visited Dec. 17, 2008).
66. See Parents Involved, 127 S. Ct. at 2768 (schools may not discriminate against whites in school admissions); Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 227 (1995). See also Holning Lau, Formalism: From Racial Integration to Same-Sex Marriage, 59 HASTINGS L.J. 843, 845-46 (2008) (arguing that the Supreme Court’s decision in Parents Involved necessarily means that sex discrimination inherent in outlawing same-sex marriage should be judged with heightened scrutiny).
67. In re Marriage Cases, 183 P.3d 384, 443 (Cal. 2008). See also Kerrigan v. Comm’r of Pub. Health, 957 A.2d 407, 428-29 n.21 (Conn. 2008) [relative political power cannot . . . be a particularly weighty factor [in determining the level of scrutiny under the Equal Protection Clause], let alone a controlling one. For example, it cannot be said that males, as a group, have been relegated to such a position of political powerlessness as to require special judicial protection. Nonetheless, laws differentiating between the sexes which disadvantage males as well as females . . . must be subjected to heightened scrutiny . . . . current political powerlessness cannot be a requirement for recognition as a protected class because other groups, including African-Americans . . ., continue to be accorded . . . protection even though they are not lacking in political power. (citations omitted).
68. Although there are often differences between the average or statistical ability or preferences of men and women, there is usually significant overlap between the sexes for any characteristic. See Attorney Gen. v. Mass. Interscholastic Athletic Ass’n, Inc., 393 N.E.2d 284, 293 (Mass. 1979) (stating that there is large variation in physical characteristics and ability among boys and girls); Force v. Pierce City R-VI Sch. Dist., 570 F. Supp. 1020, 1028-29 (W.D. Mo., 1983) (finding that sex-based classifications cannot predict how good a given individual is in a particular sport). Sex cannot predict which sex a person will be attracted to, since between 1-10% of people are non-heterosexual. See infra note 139. Additionally, while most men are taller, heavier, and/or stronger than most women, some women are taller, heavier, and/or stronger than some men. See
characteristic that most people do not choose and have no influence over. Therefore, using sex or race to determine legal rights and burdens violates "the basic concept of our system that legal burdens should bear some relationship to individual responsibility." Additionally, sex has long been used to confer disadvantage upon people and as a basis for discrimination. However, only race, not sex, is protected by the highest scrutiny.

Several states protect their citizens from sex discrimination under their state constitution using the strict scrutiny standard. The California Supreme Court has explained why sex should be judged with strict scrutiny under the equal protection clauses of both its constitution and the United States Constitution:

Sex, like race and lineage, is an immutable trait, a status into which the class members are locked by the accident of birth. What differentiates sex from nonsuspect statuses, such as intelligence or physical disability, and aligns it with the recognized suspect classifications is that the characteristic frequently bears no relation to ability to perform or contribute to society.

II. SEX-BASED CLASSIFICATIONS CAUSE NEGATIVE EFFECTS

The government may not classify based on race if a race-neutral way of classifying can be used. Similarly, the government should not be permitted to classify based on sex if a sex-neutral way of classifying can be used. There is almost always a way to describe a group of people other than by their sex, and such a description will usually more accurately reflect the aims of the classification than a description based on sex. For example, the government

Sylvia A. Law, Rethinking Sex and the Constitution, 132 U. PA. L. REV. 955, 1007 (1984) ("The problem posed by laws that classify explicitly on the basis of sex is that sex, as a proxy for some functional characteristic, is often inaccurate in relation to particular individuals.").

69. Although some people opt to change their sex through surgery. See Conway, supra note 57.


72. See, e.g., Catholic Charities of Sacramento, Inc. v. Superior Court, 85 P.3d 67, 93 (Cal. 2004) (holding that California uses strict scrutiny to evaluate sex discrimination); Guard v. Jackson, 940 P.2d 642, 643-44 (Wash. 1997) (en banc) (holding that sex classifications are subject to strict scrutiny under the Washington constitution); Baehr v. Lewin, 852 P.2d 44, 67 (Haw. 1993) (holding that sex classifications are subject to strict scrutiny under the Hawaii constitution).

73. Sail'er Inn, Inc. 485 P.2d at 540. See also Catholic Charities, 85 P.3d at 93 (reaffirming that California uses strict scrutiny to evaluate sex discrimination).


75. See Nguyen v. Immigration and Naturalization Service, 533 U.S. 53, 82 (2001) (O'Connor, J., dissenting) ("The existence of comparable or superior sex-neutral alternatives has been a powerful reason to reject a sex-based classification.").

76. See, e.g., Attorney Gen. v. Mass. Interscholastic Athletic Ass’n, 393 N.E.2d 284, 293 (Mass. 1979) (explaining that sex-neutral standards such as height, weight, or skill can be used in
may decide that it needs people available in the event of a military emergency, but they must be strong. It should require a strength test, rather than assuming that all women will not be able to fulfill the role and all men will be able to.77 Indeed, contrary to the assertion that treating the sexes differently is justified due to the “unique physical characteristics of the sexes,”78 there is little, if anything, that applies to all men and not to any women and vice versa.79

Sex-based classifications, like race-based classifications, demean the dignity of individuals because they are inherently unfair.80 Additionally, sex-based classifications reinforce sex-based stereotypes and notions of difference,81 and harm those who do not conform to sex-based stereotypes.82 Therefore, any sex-based classification should be held unconstitutional unless it meets the standard that applies to race: it is narrowly tailored to meet a compelling governmental aim and there is no sex-neutral way to accomplish the same aim.83

A. There is inherent indignity and injustice in being classified by sex

There is inherent indignity and injustice in being classified by sex. The classification is based on unchangeable physical characteristics instead of individual abilities and preferences. In classifying people based on sex, the government in a real sense looks at a person’s naked body and decides what rights that person has based upon judgments of that naked body. This act of declothing strips people of their dignity.84 The United States Supreme Court has

place of sex to determine who may play on a particular sports team). See also infra Part III:B:2 and notes 224-25 and accompanying text for sex-neutral classification alternatives in sports.


79. See supra note 68 and accompanying text. Even things that might immediately jump to mind such as “all men have penises,” or, “all women have vaginas” are not strictly true. See Kastl v. Maricopa County Cnty. College Dist., No. Civ. 02-1531, 2004 U.S. Dist. LEXIS 29825, *8 (D. Ariz. 2004) (unpublished opinion) (stating that “neither a woman with male genitalia nor a man with stereotypically female anatomy” may be denied an employment benefit). See generally, AMY BLOOM, NORMAL: TRANSSEXUAL CEOS, CROSSDRESSING COPS, AND HERMAPHRODITES WITH ATTITUDE 36 (Vintage Books 2002) (explaining that some men have vaginas). Even chromosomally, men and women are not uniform. See In re Estate of Gardiner, 42 P.3d 120, 133 (Kan. 2002) (explaining that not everyone has XX or XY chromosomes). Additionally, a pregnant man recently gave birth to a baby girl. See Guy Trebay, He’s Pregnant. You’re Speechless., N.Y. TIMES, June 22, 2008, at ST1.

80. See infra Part II:A.

81. See infra Part II:B.

82. See infra Part II:C.

83. See Frontiero v. Richardson, 411 U.S. 677, 688 (1972) (plurality opinion) (stating that because of their similarities to race classifications, sex classifications should be examined under the strict scrutiny standard); City of Richmond v. J.A. Croson Co., 488 U.S. 469, 507 (1989) (holding that race-neutral alternatives must be examined and used whenever possible as an alternative to race based classifications).

84. Cf. Kelley v. Johnson, 425 U.S. 238, 251 (1976) (Marshall, J., dissenting) (stating that the Constitution protects the “values of privacy, self-identity, autonomy, and personal integrity”). This conception of the Constitution should include protection from an invasive bodily look by the state.
stated that a principal reason that race is a forbidden classification is that it “demeans the dignity and worth of a person to be judged by ancestry instead of by his or her own merit and essential qualities.”\textsuperscript{85} Similarly, it demeans the dignity and worth of a person to be judged and accorded rights based on sex or sex stereotypes instead of his\textsuperscript{86} own individual preferences, capacities, essential qualities, and merit.\textsuperscript{87}

A girl who is not allowed to play on a boys’ basketball team for which she is qualified, simply because of her sex, is demeaned and treated unfairly because sex has nothing to do with the game of basketball, whereas the ability to pass, dribble and shoot do.\textsuperscript{88} Likewise Sam, a female, wanting to marry Julie, is similarly situated to a male, also wanting to marry Julie. It is fundamentally unfair and demeaning to deny Sam the right to marry Julie just because of the shape of her naked body, \textit{i.e.}, her sex. Likewise, it is demeaning for a person who does not conform to sex-stereotypes in appearance to be pressured to conform to a sex-stereotypical appearance because of the daily occurrences of people questioning her right to enter a particular bathroom, or not knowing what pronoun to address her by.\textsuperscript{89}

Using race or sex to pre-judge an individual may be efficient if there are average differences between groups.\textsuperscript{90} However, doing so violates the “meritocratic ideal” that rights and burdens should be determined by individual merit and ability, instead of on characteristics that may predict a general trend, such as race or sex.\textsuperscript{91} Thus, permitting sex classifications violates the basic principle that similarly situated people should be treated the same way.\textsuperscript{92} Only if

\begin{itemize}
  \item Most people would acknowledge that only one’s doctor, one’s chosen partner, and in the case of children, one’s parents, should have the right to gaze upon one’s naked body. The State should not be added to this list.
  \item \textsuperscript{85} Rice v. Cayetano, 528 U.S. 495, 517 (2000).
  \item \textsuperscript{86} One of the consequences of eliminating sex-based classifications, recommended in Part III:D, is the use of sex-neutral pronouns. Therefore, to be consistent with this recommendation, this Article will from time to time use the sex-neutral pronouns “se” (pronounced like “see”) instead of “he” or “she,” and “hir” (pronounced like “here”) instead of “her,” or “him.” \textit{Cf.} Leslie Feinberg, \textit{We Are All Works In Progress, in This Is What Lesbian Looks Like: Dyke Activists Take On the 21ST Century} 81 (Kris Kleindienst, ed., Firebrand Books, 1999) (discussing gender neutral pronouns).
  \item \textsuperscript{87} \textit{Frontiero}, 411 U.S. at 686-87 (acknowledging that sex stereotypes have often relegated females to inferior legal status without regard to their individual characteristics).
  \item \textsuperscript{88} See O’Connor v. Bd. of Educ., 449 U.S. 1301, 1307-08 (1980) (upholding school regulation that refused to allow a girl to try out for the boys’ basketball team even though her playing ability was equal to the boys on the team).
  \item \textsuperscript{89} \textit{See infra} Part II:B:3-4.
  \item \textsuperscript{90} See Leticia M. Saucedo, \textit{The Employer Preference for the Subservient Worker and the Making of the Brown Collar Workplace}, 67 OHIO ST. L.J. 961, 988 (2006) (explaining that discrimination and stereotypes are often efficient because they are based on sufficiently accurate generalizations).
  \item \textsuperscript{91} See Katherine M. Franke, \textit{What’s Wrong with Sexual Harassment?}, 49 STAN. L. REV. 691, 705 (1997) (social benefits should be based on merit, not sex).
  \item \textsuperscript{92} \textit{Id.} at 706.
\end{itemize}
all men were different from all women in a relevant way would a classification based on sex be justified. Sex-based classifications that ascribe a characteristic to an individual based on sex, when not all individuals of that sex have the characteristic, are inherently unfair and insulting.

The following cases illustrate the type of injustice that occurs when people are afforded different rights and burdens based on their sex. In *Michael M. v. Superior Court*, the United States Supreme Court upheld a statute which punished the male but not the female partner in teenage sexual intercourse. This was inherently unfair because similarly situated teenagers having consensual sexual intercourse were treated differently under the law. The male partner was criminally convicted, while the female partner was not. In *Nguyen v. Immigration and Naturalization Service*, the Court upheld a statute requiring U.S. citizen fathers, but not U.S. citizen mothers, to register with a court in order to acquire United States citizenship for their foreign-born, out-of-wedlock children. Nguyen, a man who had raised and developed a relationship with his child for two decades was told by the Court that in its opinion, most unmarried men do not raise their children. Therefore, the Court upheld a law that allowed Nguyen’s child to be deported because Nguyen had failed to register his relationship with his child. In essence, the Court was not looking at Nguyen as an individual and judging him according to his actual behavior. Instead, it was substituting a sex stereotype in the place of the individual, and in the process, demeaning and denying the man’s lived experience.

When the government judges a person according to sex and accordingly metes out legal rights, obligations, and burdens, it judges according to factors that are not within a person’s control, factors that go to the very root of what is

93. *See* Attorney Gen. v. Mass. Interscholastic Athletic Ass’n, 393 N.E.2d 284, 293 (Mass. 1979) (recognizing that “if all high school boys outstripped all high school girls in all athletic endeavors, total separation might be justifiable”).

94. Force v. Pierce City R-VI School Dist., 570 F. Supp. 1020, 1029 (W.D. Mo. 1983) (“[G]ender based classification which results from ascribing a particular trait or quality to one sex, when not all share that trait or quality, is... inherently unfair.”).

95. *See* Craig v. Boren 429 U.S. 190, 214 (1976) (Stevens, J., concurring) (explaining that even if more young men abuse alcohol than young women, it is an insult to the young men who do not abuse alcohol to ascribe the abusing characteristic to them).


97. *Id.* at 468.


99. *Id.* at 66-67.

100. *Id.* at 57.

101. *Id.* at 56-57.

102. Although people can change their sex through surgery, surely such surgery should not be required before people receive equal rights and obligations under the law. *See* Varnum v. Brien, No. 07-1499, 2009 Iowa Sup. LEXIS 31, at *76 (Iowa Apr. 3, 2009) (finding that “the immutability prong of the suspectness inquiry surely is satisfied when the identifying trait is so central to a person’s identity that it would be abhorrent for government to penalize a person for refusing to change it”) (citations omitted). *Cf.* In re Marriage Cases, 183 P.3d 384, 442 (Cal. 2008)
private. The United States Supreme Court in Lawrence v. Texas, found that people must be protected from government gaze upon the private consensual sexual conduct of adults. So too, should people be protected from government gaze upon the private nature of their naked bodies. Just as it demeans a homosexual person's existence to be judged according to private sexual behavior, so too, does it demean an individual to be judged according to private sex characteristics. Only the strongest state interest, such as the need to protect human life, should intrude on a person's dignity and de-clothe him in the face of the law.

B. Sex-based classifications reinforce sex-based stereotypes

The United States Supreme Court has said that state actors should not rely on "fixed notions concerning the roles and abilities of males and females." Contrary to this ideal, affording different rights based on sex relies on fixed ideas of the roles and abilities of males and females. In the process of classifying people by sex, the government reinforces sex stereotypes.

In Nguyen v. Immigration and Naturalization Service, the United States Supreme Court reinforced and enforced its stereotype that women, not men, bear the burden of care for their children. Because Nguyen was a man and not a

(people should not be required to change their sexual orientation before receiving equal protection under the law).

103. A person's genitals are usually kept extremely private.
105. Id.
106. Id. Even though it is often easy to tell what sex a person is, this identification is usually aided if the person conforms to sex stereotypes in hair length and dress. It is often harder to tell the sex of people who do not conform to such stereotypes. See infra notes 126-127 and accompanying text.
107. See supra note 28 and accompanying text.
109. See Miss. Univ. for Women, 458 U.S. at 730 (policy of excluding men from nursing school was illegitimate because it made "the assumption that nursing is a field for women a self-fulfilling prophecy."); Orr v. Orr, 440 U.S. 268, 283 (1979) ("Legislative classifications which distribute benefits and burdens on the basis of gender carry the inherent risk of reinforcing the stereotypes about the 'proper place' of women and their need for special protection."); Force v. Pierce City R-VI School Dist., 570 F. Supp. 1020, 1029 (W.D. Mo. 1983) ("[G]ender based classification which results from ascribing a particular trait or quality to one sex ... tends only to perpetuate 'stereotypic notions' regarding proper roles of men and women."). See also Robinson, supra note 40, at 323 (sex segregation in sports increases sex stereotyping); National Organization for Women Opposes Single Sex Public Education as "Separate and Unequal" (Oct. 24, 2006), available at http://www.now.org/press/10-06/10-24.html ("[S]ex segregation is likely to increase sex stereotyping."). (last visited Dec. 17, 2008); cf. Kerrigan v. Comm'r of Pub. Health, 957 A.2d 407, 419 (Conn. 2008) ("The very existence of the classification gives credence to the perception that separate treatment is warranted.").
woman, his child was deported to a country halfway across the globe. 111 Nguyen no longer had the opportunity to participate in his child’s life, thus reinforcing and perpetuating the Court’s stereotype that women, not men, develop meaningful relationships with their children.

1. The same-sex marriage restriction reinforces sex stereotypes

The opposite-sex requirement for marriage reinforces and relies upon ideas of what is an appropriate sex role for a person. The ban on same-sex marriage teaches society that a woman’s role is to marry a man and man’s role is to marry a woman. 112 This classification reinforces stereotypes that assume that men and women have different roles, preferences, and desires. 113 Further, in a culture where family is defined by marriage and blood relationships, this classification sends the message that a man cannot set up a real family with a man, whereas he could with a woman, and vice versa. For example, a child said to one of her lesbian aunts when they were finally able to legally marry, “now you’re really my auntie.” 114

2. Sex segregation in sports reinforces sex stereotypes

The law 115 allows a sport to be offered only to boys at a school (for example, football), or to only girls (typically, softball). This reinforces cultural judgments about which sports are masculine and which are feminine, and what characteristics boys have as opposed to girls. 116 Additionally, men often receive

111. See id. at 57-58.
113. See Varnum v. Brien, No. CV5965, 2007 WL 2468667 (Iowa Dist.) (Trial Order) (finding that “[s]ex-role conformity remains embedded in Iowa marriage law. . . . [M]ale Plaintiffs must conform to the State's view that men should fall in love with, be intimate with and marry only women, while female Plaintiffs must conform to the State's view that women should fall in love with, be intimate with and marry only men. In fact, these are old and overbroad stereotypes that do not reflect the diversity of individual men and women”); Widiss, supra note 112, at 463; Frontiero v. Richardson, 411 U.S. 677, 684-85 (1973) (plurality opinion) (describing the long history of sex stereotypes).
114. In re Marriage Cases, 49 Cal. Rptr.3d 675, 760 n.22 (Cal. Ct. App. 2006) (Kline, J., dissenting). See also Kerrigan, 957 A.2d at 475 n.77 (describing how children of same sex couples who have to explain to their friends why their parents cannot marry feel like their family is “less valid” than others); N.J. Civil Union Review Com., First Interim Rep., 11-12 (Feb. 19, 2008), available at http://www.nj.gov/oag/dcr/downloads/1st-InterimReport-CURC.pdf (last visited Dec. 17, 2008) (describing children’s impressions when gay couples are not allowed to marry).
115. 34 C.F.R. § 106.41(b) (2008) (allowing schools to operate sex segregated teams); see infra Part III:B.
116. Deborah Brake, The Struggle for Sex Equality in Sport and the Theory Behind Title IX, 34 U. MICH. J.L. REFORM 13, 140-41 (2001) (explaining that sports viewed as masculine are football,
more benefits, such as greater prestige and publicity, from being on a men’s sports team, than women receive from being on women’s sports team, even when the sports are the same. This reinforces the stereotype that athletics is properly a man’s space. Sex segregation in sports also reinforces stereotypes such as: men are better at physical activities than women and a man who cannot exhibit higher athletic ability than a woman is not a real man.118

It is true that some schools and sports teams allow children of both sexes to play on the same team, such as allowing girls to play on the football team.119 Some schools provide non-sex-stereotypical teams for their students, such as girls’ wrestling or ice hockey teams.120 However, the applicable law, Title IX, allows schools to put women and girls into less aggressive sports, while emphasizing those sports for boys and men.121 Thus, the law allows a school to keep girls from playing “contact sports” such as football, hockey, and basketball.122 Because the law allows this, many schools follow societal stereotypes and offer sports according to sex stereotypes.123 By keeping girls out of “warrior”-type sports, this practice reinforces core cultural beliefs about immutable differences between males and females.124

3. Sex segregation in bathrooms reinforces sex stereotypes

Sex-segregated bathrooms are a daily enforcer of sex differences.125 If a person does not conform to sex-stereotypes in appearance, he126 will be harassed

---

117. See Robinson, supra note 40, at 339, 354.
118. Because separate sports teams reinforce the stereotype that men are better at sports than women, a man who is not good at sports, or is worse than an individual woman is often made fun of. For example, he is likely to be told by his friends, “Don’t be such a girl,” or told that he “throws like a girl.” See Pat Griffin, Homophobia in Sport: Addressing the Needs of Lesbian and Gay High School Athletes, in THE GAY TEEN: EDUCATIONAL PRACTICE AND THEORY FOR LESBIAN, GAY, AND BISEXUAL ADOLESCENTS 53, 56 (Gerald Unks, ed., 1995) (“Studies of locker-room talk among men describe consistent patterns of antiewomen and antigay interactions.”).
120. Id.
121. See Brake, supra note 116, at 138.
122. See 34 C.F.R. § 106.41(b) (2008) (allowing schools to operate separate teams for boys and girls and exclude a person from trying out for a “contact sport” team designated for the other sex); Mercer v. Duke Univ., 401 F.3d 199 (4th Cir. 2005) (recognizing that a school may exclude a girl from a boys’ football team under the contact sport exemption).
123. See Brake, supra note 116 at 141, 145.
124. Suzanne Sangree, Title IX and the Contact Sports Exemption: Gender Stereotypes in a Civil Rights Statute, 32 CONN. L. REV. 381, 384 (2000). One of these core beliefs is the stereotype that girls and women are inherently weak.
125. Even though the law does not require the maintenance of separate-sex bathrooms, it upholds the common practice of maintaining separate-sex bathrooms and excluding people based on sex from one bathroom or the other. See, e.g., Goins v. West Group, 635 N.W.2d 717, 723 (Minn. 2001) (stating that having sex segregated bathrooms is a traditional and accepted practice); Norwood v. Dale Maint. Sys., Inc., 590 F. Supp. 1410, 1421 (N.D. Ill. 1984) (upholding the
and questioned upon entering a bathroom in which there are people who think sex does not belong, encouraging him to conform to sex stereotypes. For example, both my sister and I have had multiple experiences of walking into a women’s bathroom, and being told “Hey, you’re in the wrong place!” or “Excuse me, this is the women’s bathroom.” These incidents happened when we had short hair, and so it was more difficult for people to tell our sex simply by looking at us. As a result, I felt pressure to conform to female stereotypes and grow my hair. These are not uncommon incidents for people who do not conform to how their sex is “supposed” to look. Daily segregation, even in the realm of bathrooms, reinforces stereotypes such as: men and women are different in important ways; men and women should not share a bathroom together because of these differences; men and women should look different so no one is confused about who belongs in what bathroom.

4. Sex-specific pronouns reinforce sex stereotypes

Use of the sex-specific pronouns “he” or “she” unconsciously makes it frequent and normal to classify a person by sex, even if the person’s sex has nothing to do with the point of the conversation, reference, or thought. This reinforces stereotypical notions of fundamental difference between men and women. The word “he,” when used, produces images and ideas of males, and the word “she,” when used, produces images and ideas of females. Therefore, attaching the word “he” or “she” to a person does two things: 1) it causes others to expect that person to act in a way that conforms to their expectations of male or female; and 2) it causes that person to be pressured to conform to sex stereotypes. For example, one of the first times I was mistaken for male, by
a librarian who said to a colleague, “He is looking for a book,” I felt myself automatically stand up taller. I felt stronger and bigger than I usually feel. I was subconsciously aligning my self-image with all the stereotypes I associated with being a “he,” mentally accumulated over years of living in a gendered and sex-segregated society.

Sex-based stereotyping, like other types of stereotyping and discrimination, has long been recognized as restricting the ability of people to fully participate in society and develop their individual talents and capabilities. Sex stereotypes force some people to act against their inherent abilities and talents in order to fit into sex roles that have been defined by the culture around them. Thus, sex stereotypes limit opportunities by making it difficult to discover and develop talents and abilities. For example, if a school allows girls but not boys to play tennis as an after-school sport, a boy who otherwise may have gone on to be a world-class tennis player may never discover his tennis talent.

C. Sex-based classifications exacerbate discrimination against and otherwise cause harm to people who do not conform to sex stereotypes

In addition to being unfair and unjustified, sex-based classifications also alienate and exacerbate discrimination against intersex, transgender, homosexual and other gender non-conforming people. People who do not conform to sex-stereotypes have faced and continue to face longstanding and harmful discrimination and hatred in society. In the middle ages in Europe, intersexed people, those born with ambiguous genitalia, were compelled to choose a sex...

132. See, e.g., United States v. Virginia, 518 U.S. 515, 532 (1996) (stating that denying women equal status with men denies them opportunity to participate in and contribute to society); Frontiero v. Richardson, 411 U.S. 677, 684 (1973) (plurality opinion) (stating that there is a long history of sex discrimination, which has hurt women); Sangree, supra note 124, at 437 (explaining that the law, by allowing schools to offer contact sports exclusively to men, has denied women “the opportunity to try the sport[s] and to develop proficiency at [them]”).

133. See Frontiero, 411 U.S. at 684 (stating that sex discrimination, while based on “romantic paternalism” in fact kept “women, not on a pedestal, but in a cage”); Alice H. Eagly, Sex Differences in Social Behavior: Comparing Social Role Theory and Evolutionary Psychology, 52 AM. PSYCHOLOGIST 1380, 1381 (1997) (explaining that gender-role expectations pressure people to conform to what they might otherwise not choose to do); Jessica Knouse, Using Postmodern Feminist Legal Theory to Interrupt the Reinscription of Sex Stereotypes Through the Institution of Marriage, 16 HASTINGS WOMEN’S L.J. 159, 162 (2005) (recognizing that cultural expectations impact desires and behavior).


135. Experts estimate that about one in 2,000 people is born with ambiguous genitalia, and so, at birth, does not fit into cultural ideal of either male or female. Intersex Society of North America, http://www.isna.org/tag/frequency (last visited Dec. 14, 2008). The United States has slightly over 300,000,000 people. The World Factbook, https://www.cia.gov/library/publications/the-world-factbook/geos/us.html (last visited Dec. 14, 2008). Therefore, more than 150,000 people in the United States are intersexed (300,000,000 ÷ 2000 = 150,000).
to identify with or face death. Today they are often forced to undergo unnecessary and damaging surgery at birth, simply so they can fit into a “male” or “female” ideal. Like racial minorities, transgender and non-heterosexual people have faced violence and widespread discrimination in employment, housing, and the law, from colonial times to the present.


137. Id. at 23-24.

138. Transgender people are people “whose appearance, personal characteristics or behaviors differ from stereotypes about how men and women are ‘supposed’ to be.” National Gay and Lesbian Task Force, http://www.thetaskforce.org/reports_and_research/trans_equality (last visited Dec. 14, 2008). This group includes transsexuals, or people who were born one sex but identify as the other sex. Estimates are that the prevalence of male to female transsexualism is one in 2,500 in the United States. Conway, supra note 57. Because the United States has over 300,000,000 people, see supra note 135, at least 120,000 people do not fit neatly into the box of either “male” or “female,” because they are male to female transsexuals.

139. This group includes lesbians, bisexuals, and gay men, who do not fit into the sex stereotype of what sex they should be sexually attracted to. The rate of people who are non-heterosexual is estimated as being between 1%-10%. See, e.g., Kerrigan v. Comm'r of Pub. Health, 957 A.2d 407, 440 n.30 (Conn. 2008) (collecting sources that conclude the percentage is between 2-13%); Richard C. Pillard and James D. Weinrich, Evidence of Familial Nature of Male Homosexuality, 43 ARCH. GEN. PSYCHIATRY 808 (1986) (background rate of white male homosexuals is 4%); ROBERT T. MICHAEL ET AL., SEX IN AMERICA: A DEFINITIVE SURVEY 176 (Warner Books, 1994) (incidence of female homosexuality is 1.4%); ALFRED KINSEY, SEXUAL BEHAVIOR AND THE HUMAN MALE 623 (W.B. Saunders Co. 1948) (data indicates that at least 37% of the male population has engaged in homosexual behavior between adolescence and old age).

140. 15.3% of hate-crimes are motivated by “sexual-orientation bias.” U.S. Department of Justice, FBI, Hate Crimes Statistics 2006, http://www.fbi.gov/ucr/hc2006/victims.html (last visited February 20, 2009). See also, Farmer v. Brennan, 511 U.S. 825, 849 (1994) (recognizing that a transsexual will be subject to a heightened risk of violence in a prison environment compared to other prisoners); Tates v. Blanas, No. CITV S-00-2539 OMP P, 2003 U.S. Dist. LEXIS 26029, at *27 (E.D. Cal. March 6, 2003) (stating that in some cases prison officials have a duty to treat transsexuals differently in order to protect them from violence); Chambers v. Babitt, 145 F. Supp. 2d 1068, 1073 (D. Minn. 2001) (stating that LGBT youth are subject to a heightened risk of violence); William B. Rubenstein, The Real Story of U.S. Hate Crimes Statistics: An Empirical Analysis, 78 TUL. L. REV. 1213, 1222 (2004) (finding that LGBT people have high rates of hate crimes committed against them). Transgender people are subject to increased violence and lack of protection from the legal system. LESLIE FEINBERG, TRANS LIBERATION: BEYOND PINK OR BLUE 10-11 (Beacon Press 1998) (because of social punishments such as police brutality, gang rape, and loss of child visitation rights, “[t]rans people are still literally social outlaws”).

141. See Lawrence v. Texas, 539 U.S. 558, 602 (2003) (Scalia, J., dissenting) (“Many Americans do not want persons who openly engage in homosexual conduct as partners in their business, as scoutmasters for their children, as teachers in their children's schools, or as boarders in their home.”); Ulane v. E. Airlines, Inc., 742 F.2d 1081, 1081-82 (7th Cir. 1984) (holding employment dismissal of a male-to-female transsexual for being a transsexual did not violate Title VII).

142. See Lawrence, 539 U.S. at 602 (Scalia, J., dissenting) (“Many Americans do not want persons who openly engage in homosexual conduct . . . as boarders in their home.”).

Finally, others who do not fit into sex-stereotypes, such as “sissy-boys,” are often subjected to humiliation and bullying. Similar to “White Supremacy,” which was maintained by racial classifications, sex classifications perpetuate a “Gender Normal Supremacy” where people who conform to gender stereotypes are deemed superior to those who do not, such as tomboys, “sissies,” and homosexuals.

The more sex stereotypes are reinforced and affirmed in society, the more difficult it is for people, whose natural self-expression is in conflict with sex-role expectations, to accept themselves and to function well within society. Constant societal messages that they are strange and different causes reduced self-esteem and achievement, as well as increased incidents of depression, suicide, harassment, and hate crimes.

Having the government augment the separation of the sexes and reinforce sex-stereotypes, by giving different rights to people based on their sex, exacerbates the problems associated with stereotyping. The Supreme Court constitutional amendment that had been motivated solely to make homosexuals “unequal to everyone else”); In re Estate of Gardiner, 42 P.3d 120, 135 (Kan. 2002) (stating that a male-to-female transsexual cannot qualify, in Kansas, as a woman for the purposes of marriage); TEX. PENAL CODE ANN. § 21.06 (2006) (making homosexual sex a crime) (invalidated by Lawrence v. Texas, 539 U.S. 558 (2003)); Darren Lenard Hutchinson, Dissecting Axes of Subordinated: The Need for a Structural Analysis, 11 AM. U. J. GENDER SOC. POL’Y & L. 13, 13-14 (2002) (stereotypes and prejudices can disadvantage LGBT people in criminal trials).

144. See Kerrigan, 957 A.2d at 432 (“Gay persons have been subjected to and stigmatized by a long history of purposeful and invidious discrimination that continues to manifest itself in society.”); High Tech Gays v. Def. Indus. Sec. Clearance Office, 668 F. Supp. 1361, 1369 (N.D. Cal. 1987) (“Lesbians and gay men have been the object of some of the deepest prejudice and hatred in American society.”); Harris M. Miller II, Note, An Argument for the Application of Equal Protection Heightened Scrutiny to Classifications Based on Homosexuality, 57 S. CAL. L. REV. 797, 824-25 (1984) (describing history of LGBT discrimination).


146. See Kerrigan, 957 A.2d at 433 (noting that “anti-homosexual attitudes are associated with psychological distress for homosexual persons and may have a negative impact on mental health, including a greater incidence of depression and suicide”) (citations omitted); Kimberly F. Balsam et al., Mental Health of Lesbian, Gay, Bisexual, and Heterosexual Siblings: Effects of Gender, Sexual Orientation, and Family, 114 J. ABNORMAL PSYCHOL. 471, 474 (2005) (finding that nonheterosexual people have higher rates of suicide attempts and more anxiety and depression than heterosexual people); Anthony R. D’Augelli, Developmental Implications of Victimization of Lesbian, Gay, and Bisexual Youths, in STIGMA AND SEXUAL ORIENTATION: UNDERSTANDING PREJUDICE AGAINST LESBIANS, GAY MEN, AND BISEXUALS 187, 201-02 (Gregory Herek, ed., 1998) (explaining that lesbian, gay, bisexual and transgender youth are more likely to attempt suicide than heterosexual youth, and 30% have attempted suicide).

149. See Kerrigan, 957 A.2d at 433.

150. See American Psychological Association, http://www.apa.org/releases/gaymarriage.html (July. 28, 2004) (“Denying same-sex couples legal access to civil marriage is discriminatory and
recognized that when a law makes homosexual conduct illegal, that law “is an invitation to subject homosexual persons to discrimination both in the public and private spheres.”151 Similarly, when the law mandates different treatment for people based on sex and sex stereotypes, people tend to disapprove of those whose sex is not obvious or who do not conform to sex stereotypes.152

1. Outlawing same-sex marriage harms LGBT people

Lesbian, gay, bisexual, and transgender (LGBT) people often cannot legally marry their partners, unlike similarly situated heterosexual people. This causes society to view LGBT people as “other,” relegated to the outside by law.153 It only exacerbates prevailing sex stereotypes for the government to, in effect, say to its people: “Two men are not equal in the eyes of the law and will not receive the same rights as a man and a woman, and vice versa.” Under this rule of law, teenagers encountering a same-sex couple walking down the street with their arms over each others’ shoulders may feel state-sanctioned in throwing ill-intentioned, sarcastic, and disparaging remarks the couple’s way.154 Even worse, state-sanctioned discrimination against those who do not conform to sex stereotypes encourages hatred and hate crimes, such as the killing of gay university student Matthew Shepard in 1998.155 Finally, outlawing same-sex marriage harms the children of same-sex couples by denying them many legal benefits and protections.156

151. Lawrence v. Texas, 539 U.S. 558, 575 (2003). See also Kerrigan, 957 A.2d at 419 (“The very existence of the classification gives credence to the perception that separate treatment is warranted.”).

152. Cf. Planned Parenthood of Southeastern Pa. v. Casey, 505 U.S. 833, 995 (1992) (Scalia, J., concurring and dissenting) (explaining that if the law indicates approval or disapproval of a certain behavior, people will tend to believe the behavior is good or bad, in accordance with the law); Kerrigan, 957 A.2d at 419 (“The very existence of the classification gives credence to the perception that separate treatment is warranted.”).

153. See id.; N.J. Civil Union Review, supra note 114, at 11-12.

154. I have witnessed this type of incident, and have met many LGBT people who will not hold their partner’s hand in public for fear of negative comments or violence. Hate crimes based on sexual orientation comprise 15.3% of all hate crimes. See supra note 140.

155. See supra note 140; Gerard Wright, Gay Grief in Cowboy County; Matthew Shepard was Lashed to this Fence and Pistol-whipped. He Died Five Days Later. His Crime? He was a Homosexual in the American West, THE GUARDIAN (London), Mar. 27, 1999, available at http://www.guardian.co.uk/books/1999/mar/27/books.guardianreview7; cf. Lawrence v. Texas, 539 U.S. 558, 575 (2003) (“When homosexual conduct is made criminal by the law of the state, that declaration in and of itself is an invitation to subject homosexual persons to discrimination both in the public and in the private spheres.”).

156. See infra Part III.A.3.
2. Sex-segregated sports harm “sissy” boys and LGBT people

When children are separated based on sex for school athletic programs, traditional and stereotypical concepts of masculinity are reinforced. Homophobia and disrespect for boys who cannot perform to high levels is exacerbated. High performing girls are unfairly excluded from teams they are otherwise qualified for.

3. Sex-segregated bathrooms harm those who fail to “look their sex”

Transgender people are not compatible with laws, or practices upheld by law, that divide people based on sex. As one state supreme court has noted, “[t]he words ‘sex,’ ‘male,’ and ‘female’ in everyday understanding do not encompass transsexuals.” Transgender people and people who do not fit into stereotypes of how people of their sex should look have a particularly difficult time when it comes to sex-segregated bathrooms. No matter which bathroom they go to, they often experience questioning, hostility, and statements that they do not belong. Transgender people may even be fired from their job for using the “wrong” bathroom.

4. Sex-specific pronouns harm transgender people

Sex-specific pronouns require us to label everyone as either male or female. Those who cannot be easily labeled, or who would like to be labeled with a different pronoun than one that people expect, are often greeted with reactions of confusion, hostility or violence.
As long as sex stereotypes are encouraged or reinforced in law, transsexuals and the intersexed, as well as others who do not conform to sex stereotypes, such as non-heterosexuals, will be rendered outcasts by law, and, like African Americans under a system of racial segregation, will have “feeling[s] of inferiority as to their status in the community that may affect their hearts and minds in a way unlikely ever to be undone.”\(^{164}\) If sex-based classifications were subjected to strict judicial scrutiny, most, if not all, sex classifications upheld by law would fall,\(^{165}\) thus ameliorating the legal and social problems of people who do not fit neatly into sex-stereotypes associated with “male” or “female.”

\[\text{III. If courts ruled that sex-based classifications must be judged with strict scrutiny, several societal changes would occur}\]

Some have argued that the Equal Rights Amendment, by explicitly mandating equality between men and women, would lead to same-sex marriage and co-ed bathrooms.\(^{166}\) However, even without the Equal Rights Amendment, same-sex marriage, co-ed bathrooms and sports teams, and sex-neutral pronouns are a constitutional imperative under the Equal Protection Clause. Under a strict scrutiny analysis, the government has the burden of showing that its classification, even one that purportedly causes no harm,\(^{167}\) is narrowly tailored to address a compelling governmental interest.\(^{168}\) As will be shown below, sex classifications in marriage rights, school sports, bathrooms, and pronouns would all fall under the strict scrutiny test.\(^{169}\) As a result, the sex stereotyping and other negative effects discussed in Part II above would be reduced.

\[\text{A. Same-sex marriage would be permitted}\]

Opposite-sex marriage requirements should be ruled unconstitutional because they treat similarly situated people differently and discriminatorily based on sex, and they are not justified by any compelling state interest.\(^{170}\) The institution of marriage is prominent in our society and is mentioned hundreds of


\(^{165}\) See infra Part III.

\(^{166}\) They then conclude that because of these possible results, such an amendment should not be added to the Constitution. See Widiss, supra note 112, at 463; Donald J. Boudreaux & A.C. Pritchard, Rewriting the Constitution: An Economic Analysis of the Constitutional Amendment Process, 62 FORDHAM L. REV. 111, 155 (1993).

\(^{167}\) This is the fallacious “separate but equal” argument. See infra Part III:A.1.


\(^{169}\) Even relying on intermediate scrutiny, these classifications should be held invalid. See supra note 31.

\(^{170}\) See supra notes 8-10 and accompanying text for why prohibiting same-sex marriage is sex discrimination.
times in statutes, affecting rights including: the ability to make medical decisions for a loved one in emergencies; rights to child and spousal support; rights of spousal immunity; and rights to tax breaks. People not able to marry their loved one often suffer emotional and legal disadvantages because they are excluded from an institution of central significance in our society. The opposite-sex marriage requirement is therefore one of the most significant examples of the negative effects government sex-based classifications can have on a person’s life.

1. “Separate but equal” treatment is not equal

Most courts addressing the issue have stated that the opposite-sex marriage requirement does not discriminate based on sex because it treats men and women equally and “each sex is . . . prohibited from precisely the same conduct,” in that each may not marry someone of the same sex. These courts look at whether a person is disadvantaged, compared to a differently-sexed person “similarly situated.” They then conclude that because a woman who cannot marry a woman is in “like circumstance” as a man who cannot marry a man, and both are disadvantaged to the same extent, there is no sex-based discrimination. Thus, this pervasive sex-based classification is justified by the theory that it treats both sexes equally, though separately; thereby upholding a “separate but equal” doctrine for sex. Under a system of race segregation, blacks


174. See, e.g., Hernandez v. Robles, 855 N.E.2d 1, 10-11 (N.Y. 2006) (finding that the opposite-sex marriage requirement is not sex discrimination because “[w]omen and men are treated alike - they are permitted to marry people of the opposite sex, but not people of their own sex.”); Andersen v. King County, 138 P.3d 963, 988 (Wash. 2006) (en banc); In re Marriage Cases, 49 Cal. Rptr. 3d 675, 706 (Cal. Ct. App. 2006) (holding that the opposite-sex marriage requirement, is a law that “merely mentions gender”, treats both sexes equally, and so is not discriminatory). But see supra note 9 and accompanying text.

and whites were treated separately and purportedly equally in that only blacks could marry blacks and go to black schools, and only whites could marry whites and go to white schools. Similarly, when same-sex marriage is not allowed, but opposite-sex marriage is, men and women are treated separately and purportedly equally in that only men may marry women, and only women may marry men.

Indeed, the Supreme Court and many state courts have ruled that any sex-based classification is only subject to the “heightened,” “intermediate scrutiny” standard under the Constitution when it “favors” one sex over the other, or treats males “worse” than similarly situated females and vice versa. This subjective standard is in contrast to race-based classifications where the Court has said, in effect, that “roughly equal” or “equal though not the same or identical” treatment is not good enough: exactly the same opportunities must be given to people, regardless of race. Just as the “separate but equal” theory does not work for race-based classifications, neither does it work for sex-based classifications. In both cases “separate but equal” laws seldom are equal in the eyes of those challenging the law. A woman wanting to marry a woman is unlikely to be comforted by being told that she may marry a man instead. A girl wanting to play on a boys’ basketball team because it is more skilled than the girls’ basketball team is unlikely to be comforted by being told that she can play

177. Heightened scrutiny refers to either “intermediate scrutiny” or “strict scrutiny.” For an explanation of the lowest level of scrutiny, “rational basis” scrutiny, see supra note 50.

178. Michael M. v. Superior Court, 450 U.S. 464, 471-73 (1981) (holding that a statutory rape law which held men but not women criminally liable for sex with an underage minor did not violate the Equal Protection Clause because both sexes bore risk; men received criminal punishment while women bore the risk of becoming pregnant); Baker, 744 A.2d at 880 n.13 (stating that opposite-sex marriage requirements are facially gender-neutral and do not have a discriminatory purpose with regard to sex, and so do not trigger equal protection analysis); Hines v. Caston Sch. Corp., 651 N.E.2d 330, 336 (Ind. Ct. App. 1995) (holding that a school dress code which prohibited boys but not girls from wearing earrings did not violate equal protection because the dress code required all students, regardless of sex, to conform to the dress code, which was based on community standards). But see Baehr, 852 P.2d at 64 (holding that strict scrutiny review should apply because the applicable statute regulates access to marital status based on sex, thus creating a sex-based classification); Westley v. Rossi, 305 F. Supp. 706, 713-14 (D. Minn. 1969) (holding that state enforcement of a public school regulation requiring boys but not girls to have short hair is constitutionally impermissible).

179. See, e.g., Loving v. Virginia, 388 U.S. 1, 12 (1967) (if a white person may marry a white person, then a black person must also be able to marry a white person, and vice versa); Brown v. Bd. of Educ., 347 U.S. 483, 494-95 (1954) (racial segregation is outlawed in schools, even if segregated schools have the same quality and educational opportunity).

180. See Varnum v. Brien, No. 07-1499, 2009 Iowa Sup. LEXIS 31, at *48 (Iowa Apr. 3, 2009) (finding that “civil marriage with a person of the opposite sex is as unappealing to a gay or lesbian person as civil marriage with a person of the same sex is to a heterosexual. Thus, the right of a gay or lesbian person under the marriage statute to enter into a civil marriage only with a person of the opposite sex is no right at all.”). Contra Andersen v. King County, 138 P.3d 963, 988 (Wash. 2006) (stating that the opposite-sex marriage requirement treats the sexes equally because although a woman may not marry a woman, she may marry a man).
on the less skilled girls’ team instead.\textsuperscript{181} A boy wanting to play on a school’s tennis team will not find it equal to be told that he can play on a football team instead.\textsuperscript{182}

Stating that both men and women have the same rights in that both may “marry someone of the opposite sex,” is a semantic trick, upholding the “separate but equal” doctrine for sex, when it was struck down decades ago in the context of race.

Indeed, people are often irreplaceable in the lives of their closest loved ones. As the Supreme Court has recognized, people are unique and not replaceable.\textsuperscript{183} In the very context of marriage, the Supreme Court acknowledged in 1967, in \textit{Loving v. Virginia}, that it constitutes a denial of fundamental rights to forbid someone from marrying the person of their choosing, even if that person could marry someone else, someone different from the one they choose.\textsuperscript{184} Therefore, if people are truly unique, then the only people “similarly situated” to Sam wanting to marry Julie are people also wanting to marry Julie, Sam’s particular and unique partner.

Furthermore, courts have recognized that the sexes are different and not fungible.\textsuperscript{185} Most people have a clear romantic preference for people of one sex over people of a different sex.\textsuperscript{186} It is not therefore equal or equivalent to tell a person se\textsuperscript{187} can marry a person of the opposite sex if se would like to marry someone of the same sex.

\textsuperscript{181.} \textit{Contra} O’Connor v. Bd. of Educ., 449 U.S. 1301, 1307-08 (1980) (upholding school regulation which refused to allow a girl to try out for the boys’ basketball team even though boys’ team was more skilled than the girls’ team).
\textsuperscript{182.} \textit{Contra} Mularadelis v. Haldane Cent. Sch. Bd., 427 N.Y.S.2d 458, 464 (N.Y. App. Div. 1980) (holding that boy may be excluded from female tennis team when school has no male tennis team but has eleven other male sports teams).
\textsuperscript{183.} See, e.g., Rice v. Cayetano, 528 U.S. 495, 517 (2000) (stating that each of us possesses a unique personality); Payne v. Tennessee, 501 U.S. 808, 832 (1991) (O’Connor, J., concurring) (each person is irreplaceable, with unique hopes, dreams and fears).
\textsuperscript{184.} \textit{Loving}, 388 U.S. at 12. \textit{See also} Perez v. Lippold, 198 P.2d 17, 25 (Cal. 1948) (striking down California’s anti-miscegenation law and finding invalid the argument that people can find someone of their own race to marry as “[h]uman beings are bereft of worth and dignity by a doctrine that would make them as interchangeable as trains”).
\textsuperscript{185.} United States v. Virginia, 518 U.S. 515, 533 (1996) (stating that the “two sexes are not fungible”); Ballard v. United States, 329 U.S. 187, 193 (1946) (“[T]he two sexes are not fungible; a community made up exclusively of one is different from a community composed of both.”); Frandsen v. County of Brevard, 800 So. 2d 757, 758 (Fla. Dist. Ct. App. 2001) (“Physical differences between men and women are enduring.”). Note that I am not disputing that the sexes may be generally different in some ways or that people are entitled to distinguish between the sexes in their personal lives; that question is up to debate, personal opinion and choice. Rather, I am arguing that the government should not distinguish between the sexes to determine legal rights and burdens.
\textsuperscript{186.} \textit{See Varnum}, 2009 Iowa Sup. LEXIS, at #48 (noting that “civil marriage with a person of the opposite sex is as unappealing to a gay or lesbian person as civil marriage with a person of the same sex is to a heterosexual”). Estimates are that, at most, 10% of the population is gay, lesbian, or bisexual. \textit{See supra} note 139.
\textsuperscript{187.} For an explanation of the sex-neutral pronouns “hir” and “se” \textit{see infra} Part III:D.
Common justifications for the opposite-sex marriage requirement include procreation and the state’s interest in the optimal rearing of children. These justifications are not narrowly tailored to meet a compelling state interest because same-sex couples do procreate and make as good parents as opposite-sex couples. Thus, if sex-based classifications were judged with strict scrutiny, the law would have to change to allow same-sex marriages.

2. Purported state interests in procreation do not meet strict scrutiny

Hawaii used to require potential marriage partners to show they were not impotent, but it changed the law, demonstrating its belief that the primary purpose of marriage was not simply to “promote and protect propagation.” The purported rationale of ensuring the propagation of the human race is outdated. The proportion of LGBT people in society is small, between one and ten percent. Thus, even if all LGBT people ceased to reproduce, no population problem would result. However, many LGBT people and same-sex couples, like many heterosexual people and opposite-sex couples, do produce and/or raise children, regardless of whether they are permitted to marry, through means that include reproductive technology and adoption.

If the state’s interest is in procreation, then it can accomplish its goal in sex-neutral ways, such as mandating that only people who intend to and are physically capable of procreation with the intended marriage partner may marry each other. This would eliminate all infertile people, including men who have had vasectomies, women who have had hysterectomies, women past menopause, as well as people with various medical conditions, from participating in marriage. However, in reality there is no legal link between the ability to marry and the ability to procreate. Instead, the children of same-sex couples are disadvantaged when their parents are denied the benefits of marriage.

188. See, e.g., Andersen v. King County, 138 P.3d 963, 969 (Wash. 2006) (stating that the legislature is rational to believe that limiting marriage to opposite-sex couples furthers “procreation, essential to the survival of the human race” as well as optimal child-rearing); Baehr v. Lewin, 852 P.2d 44, 74 (Haw. 1993) (Heen, J., dissenting) (restricting marriage to opposite-sex couples furthers the goal of “protecting the propagation of the human race”); Wilson v. Ake, 354 F. Supp. 2d 1298, 1308 (M.D. Fla. 2005) (upholding federal DOMA and state opposite-sex marriage requirements as rationally related to legitimate governmental goals, including procreation and child-rearing).


190. See Baehr v. Lewis, 852 P.2d 44, 48 n.1 (Haw. 1993) (citation omitted).

191. See id.

192. See supra note 139.


194. Goodridge v. Dep’t of Pub. Health, 798 N.E.2d 941, 961 (Mass. 2003) (stating that procreation is not the defining characteristic of a marriage, and that people who have no ability to or will not procreate are as entitled as others to marry). See also Varnum v. Brien, No. 07-1499, 2009 Iowa Sup. LEXIS 31, at * 104 (Iowa Apr. 3, 2009) (finding that the statute forbidding same-
3. Purported state interests in the best interest of the child do not meet strict scrutiny

The theory that same-sex parents make worse parents than opposite-sex parents is not supported by the scientific literature, and has been rejected by the American Psychological Association. It is likely that “best interest of the child” arguments are being used against the LGBT community, just as they have been against another unpopular group, African-American men, to “juxtapose the supposed purity of children against the supposed licentiousness of same-sex relationships.” Indeed, the best interest of children of same-sex couples is to allow their parents to marry, so the children can obtain more legal protections and long-term stability in the home.

4. Purported state interests in religion do not meet strict scrutiny

A common argument against same-sex marriage is that it goes against the teachings of various religions. However, many religious texts endorse things that are not condoned by law today. In two extreme examples, the Bible

sex marriage is “significantly under-inclusive with respect to the objective of increasing procreation because it does not include a variety of groups that do not procreate for reasons such as age, physical disability, or choice”); Kerrigan v. Comm'r of Pub. Health, 957 A.2d 407, 424-25 n.19 (Conn. 2008) (noting that there are “opposite sex couples [that] also are unable to procreate, and others [that] choose not to do so”). Karen M. Loewy, The Unconstitutionality of Excluding Same-Sex Couples From Marriage, 38 NEW ENG. L. REV. 555, 558-59 (2004) (explaining that on a marriage application one does not have to announce intention to procreate; when a person loses the ability to procreate, the person’s marriage is no less valid in the eyes of the law).

195. See infra note 199 and accompanying text; Chambers, supra note 171, at 447 (marriage provides many legal benefits for families).

196. See Varnum v. Brien, No. 07-1499, 2009 Iowa Sup. LEXIS 31, at *13-14 (Iowa Apr. 3, 2009) (noting that “[a]lmost every professional group that has studied the issue indicates children are not harmed when raised by same-sex couples, but to the contrary, benefit from them”); Hernandez v. Robles, 855 N.E.2d 1, 32 (N.Y. 2006) (Kaye, C.J., dissenting) (explaining that social science research shows that LBGT parents make as good parents as heterosexual parents); Herok, supra note 193, at 613 (explaining that studies comparing children of opposite-sex parents to children of same-sex parents have found them equally well-adjusted).


199. See Varnum, 2009 Iowa Sup. LEXIS, at *44 (noting that marriage provides same-sex couples a “stable framework within which to raise their children”); Hernandez, supra note 193 at 31-32 (Kaye, C.J., dissenting) (arguing that it is in the best interest of the children of LGBT parents to allow their parents to get married so that family decisions about property and care of sick family members can be made as a unit).

the holding of slaves and the keeping of women as property of men. 201
Additionally, as a constitutional democracy, the Constitution is the supreme law
of the land, and not a religious text which some hold as their moral law. 202
Religious institutions have never governed civil marriage. 203 They may, and
often do, make the marriage requirements for people wanting to participate in
their religious marriage ceremony narrower than is required by law. For
example, some religious institutions prohibit interdenominational, previously
divorced, or homosexual marriages within their walls. 204 Thus, religious freedom
is not jeopardized by legalizing same-sex marriage. 205 A person permitted to
marry in a town hall need not to be permitted to marry, for example, in the
Catholic Church.

The professed justifications for maintaining the opposite-sex requirement for
marriage do not stand up to a strict scrutiny analysis. 206 Instead, there are
important and compelling reasons for the elimination of the opposite-sex
marriage requirement. 207 Another area that the law would change if sex-based
classifications were judged by strict scrutiny is that of school sports.

201. See Exodus 21:7 (King James) (owning women); Leviticus 25:44 (King James) (owning
slaves).

202. See U.S. Const. art. VI (“This Constitution . . . shall be the supreme Law of the Land . . .
.”); see also Varnum, 2009 Iowa Sup. LEXIS, at *115 (noting that “civil marriage must be judged
under our constitutional standards of equal protection and not under religious doctrines or the
religious views of individuals”).

203. Andersen v. King County, 138 P.3d 963, 1033 n.8 (Wash. 2006) (en banc) (Bridge, J.,
dissenting) (explaining that religious institutions never governed civil marriage in the United
States, but they could put marriage restrictions on their own congregations, such as not allowing a
divorced person to remarried within their religious institution).

204. Id.; see The Vatican on Divorce, http://www.vatican.va/archive/ccc_css/archive/
catechism/p2s2c3a7.htm#1650 (last visited Feb. 24, 2009) (divorced people cannot marry in the
Catholic Church).

205. See Varnum v. Brien, No. 07-1499, 2009 Iowa Sup. LEXIS 31, at *114-16 (Iowa Apr. 3,
2009) (even after legalizing same-sex marriage, “[a] religious denomination can still define
marriage as a union between a man and a woman”); Kerrigan v. Comm’r of Pub. Health, 957 A.2d
407, 475 (Conn. 2008) (“Religious freedom will not be jeopardized by the marriage of same sex
couples because religious organizations that oppose same sex marriage as irreconcilable with their
beliefs will not be required to perform same sex marriages or otherwise to condone same sex
marriage or relations.”).

206. See Kerrigan, 957 A.2d at 475 (finding that religious objections to same-sex marriage do
not even meet intermediate scrutiny and as “marriage is a state sanctioned and state regulated
institution, religious objections to same sex marriage cannot play a role in our determination of
whether constitutional principles of equal protection mandate same sex marriage”). See also
disapproval of or desire to harm a particular group is not a legitimate state interest under the Equal
Protection Clause of the Constitution); Romer v. Evans, 517 U.S. 620, 634 (1996) (desire to harm
an unpopular group is not a legitimate state end).

207. See supra Part II; Michael H. v. Gerald D., 491 U.S. 110, 141 (1989) (Brennan, J.,
dissenting) (stating that in a society like ours, “liberty” must include the freedom not to conform”);
individuals of the right to choose for themselves how to conduct their intimate relationships poses a
far greater threat . . . than tolerance of nonconformity could ever do.”).
B. Sports in schools would be integrated

Current justifications for having single-sex sports at publicly-funded institutions, including schools, do not meet the proposed strict scrutiny test. Presently, publicly-funded educational institutions are permitted by law to sex-segregate in creating sports teams.208 Ironically, the statute under which this jurisprudence developed, Title IX, prohibits sex discrimination in federally-funded educational benefits.209 However, for sports teams, a “separate but equal” doctrine has developed in the regulations and court cases.210 Because Title IX requires that women have equal opportunity to play sports as men,211 it has been credited with vastly expanding the opportunities for women and girls in sports since its beginnings in the 1970s.212 Despite this significant achievement in eradicating some of the effects of historical sex discrimination, Title IX allows athletic departments to sponsor separate teams for men and women, and to permit a school to offer a sport for one sex but not the other.213 Thus, Title IX and regulations under it must change if they are truly to equalize opportunity for men and women. This is because separate facilities are “inherently unequal.”214

Under current law, if an educational institution has a men’s team and a women’s team for a particular sport, then the institution is permitted to exclude both men and women from playing on the team which is not of their sex.215 This is true even if one team plays at a higher level than the other.216 If the institution only has a men’s team for a particular sport, Title IX regulations allow girls to

208. See 34 C.F.R. § 106.41(b) (2008) (permitting schools to operate separate teams for boys and girls).
209. 20 U.S.C. § 1681(a) (2006) (“No person . . . shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving Federal financial assistance . . . .”).
210. 34 C.F.R. § 106.41(b) (2008) (allowing separate teams for boys and girls); Mercer v. Duke Univ., 401 F.3d 199, 202 (4th Cir. 2005) (permitting a school to exclude a girl from a boys’ football team); Adams v. Baker, 919 F. Supp. 1496, 1503 (D. Kan. 1996) (stating that Title IX regulations clearly contemplate separate teams for boys and girls). See also Nancy Hogshead-Makar, A Critique of Tilting the Playing Field: Schools, Sports, Sex and Title IX, 13 UCLA WOMEN’S L.J. 101, 128 n.114 (2003) (book review) (explaining that the regulations clearly contemplate separate sports teams for men and women when the sport is a contact sport and/or is based on competitive ability).
211. 34 C.F.R. § 106.41(c) (2008) (requiring “equal athletic opportunity for members of both sexes”).
212. Robinson, supra note 40, at 345.
213. 34 C.F.R. § 106.41(b) (2008) (regulation under Title IX allowing separate teams for boys and girls if the teams are based on competitive ability or the activity involved is a contact sport); Mercer, 401 F.3d at 201-02 (recognizing that Title IX did not require a school to give women the opportunity to play on a men’s football team); Force v. Pierce City R-VI School Dist., 570 F. Supp. 1020, 1024-25 (W.D. Mo. 1983) (explaining that Title IX allows schools to choose whether or not to have sex-integrated sports teams).
215. O’Connor v. Bd. of Educ., 449 U.S. 1301, 1306-08 (1980) (denying relief to a girl who wanted to play basketball on the boys’ team when there was a girls’ team, even though the girls’ team was of a lower standard than the boys’ team).
216. Id.
try out for the team, but only if it is a non-contact sport. Contact sports include “boxing, wrestling, rugby, ice hockey, football, basketball and other sports the purpose or major activity of which involves bodily contact.” If a school only has a girls’ team, or only a boys’ team, courts are split on the issue of whether a member of the excluded sex may try out for a team if it involves contact sports or if a boy is trying out for a girls’ team. Although individual schools and colleges may decide to sponsor co-ed sports, let girls play on the boys’ football or wrestling teams, or create contact sport teams for girls, they are not required to by law, and many do not.

Because sex-based classifications should be judged with strict scrutiny, Title IX and the regulations under it should be judged unconstitutional in violation of the Equal Protection Clause, insofar as they permit sex-segregation in sports teams. Football, basketball, swimming, and other sports teams at public schools and publicly-funded universities should all be co-ed. This may seem outrageous, but when there is very little physical difference between the sexes in children who have not yet reached puberty, there can be no purpose for separating such children except to reinforce notions of differences between the sexes.

217. See 34 C.F.R. § 106.41(b) (2008).
218. Id.
219. Compare Clark v. Ariz. Interscholastic Ass’n, 695 F.2d 1126, 1131 (9th Cir. 1982) (boy may be excluded from a girls’ volleyball team even if there is no boys’ volleyball team), and Mercer v. Duke Univ., 401 F.3d 199, 201-02 (4th Cir. 2005) (school may exclude a girl from a boys’ football team even when school has no girls’ football team), and State v. Hunter, 300 P.2d 455 (Or 1956) (girls may be excluded from wrestling competitions where school has no girls’ wrestling team), and Mularadelis v. Haldane Cent. Sch. Bd., 427 N.Y.S.2d 458, 464 (N.Y. App. Div. 1980) (boy may be excluded from a female tennis team at a school that had no male tennis team) with Adams v. Baker, 919 F. Supp. 1496, 1505 (D. Kan. 1996) (girl must be allowed to wrestle on boys’ wrestling team where there is no girls’ wrestling team), and Force v. Pierce City R-VI Sch. Dist., 570 F. Supp. 1020, 1031 (W.D. Mo.1983) (girl can play on a boys’ football team when there is no girls’ football team), and Gomes v. R. I. Interscholastic League, 469 F. Supp. 659, 665-66 (D.R.I. 1979) (overruled on other grounds) (boy may not be excluded from a girls’ volleyball team if there is no equivalent boys’ volleyball team).
220. See, e.g., Tamar Lewin, In Twist for High School Wrestlers, Girl Flips Boy, N.Y. TIMES, Feb. 17, 2007, at A2 (describing the lack of uniformity between schools in policies allowing girls and boys to play sports together); Robert M. Malina & Gaston Beunen, Matching of Opponents in Youth Sports, in THE CHILD AND ADOLESCENT ATHLETE 202, 204 (Oded Bar-Or ed., Blackwell Science Ltd., 1996) (stating that although co-educational sports are often offered to children under the age of twelve, after age twelve, co-educational opportunities are unavailable in the majority of cases).
221. Just as race-based integration once seemed outrageous to the majority of people. See Robinson, supra note 40, at 324-25 (stating that past race-based jurisprudence allowing racial segregation now seems outrageous, though it once seemed reasonable).
222. Except the physical difference in genital structure, but that does not impact sports ability.
223. See Robert M. Malina, Growth and Maturation Applications to Children and Adolescents in Sports, in PEDIATRIC SPORTS MEDICINE FOR PRIMARY CARE 39, 41 (Richard B. Birrer et al. eds., Lippincott Williams & Wilkins, 2002) (showing that there is little average difference in size to weight ratios between boys and girls before adolescence); id. at Appendix D (showing that there is little difference between the weight and height of boys and girls before puberty).
When the children grow older and there is actually a physical difference in the average ability\textsuperscript{224} of males and females to perform physical tasks such as running and swimming, there are sex-neutral ways of dividing teams that will be fair to both men and women. People are often divided by skill into varsity and junior varsity teams. In wrestling and rowing, people are often separated into different teams based on weight. Similarly, all sports teams can be divided based on skill, weight, strength, or other sex-neutral characteristic.\textsuperscript{225} For example, where a school used to have a boys’ team and a girls’ team, it could now have teams for people under a certain height or weight, and over a certain height or weight. Sex should not be used as a proxy for other, more relevant characteristics, simply for the purpose of “administrative ease and convenience.”\textsuperscript{226}

Common justifications for maintaining sex segregated teams could not pass a strict scrutiny they because they are not narrowly tailored to a compelling governmental interest.

1. Safety concerns do not meet strict scrutiny

Courts have stated that an objective which aims to exclude or protect members of one sex from the other sex, because they are assumed to suffer an inherent handicap or be innately inferior, is an illegitimate objective.\textsuperscript{227} Nevertheless, arguments are often made that women should be kept off men’s teams and vice versa because of safety concerns; that women will be hurt if they play sports with men because they are physically weaker and smaller than men.\textsuperscript{228} History shows that some women are able to compete with the best of the men, so clearly there is variation among women in terms of sports ability.\textsuperscript{229} If

\textsuperscript{224} There is significant overlap in size and sports ability between men and women. See supra note 68.

\textsuperscript{225} See Attorney Gen. v. Mass. Interscholastic Athletic Ass’n, 393 N.E.2d 284, 295 (Mass. 1979) (stating that sex-neutral standards such as height, weight, or skill can be used in place of sex to determine who may play on a particular team); Adams v. Baker, 919 F. Supp. 1496, 1504 (D. Kan. 1996) (stating that a school can take into account differences in size, strength, or experience, without assuming these differences are based on sex).

\textsuperscript{226} Craig v. Boren, 429 U.S. 190, 198 (1976) (recognizing that administrative ease and convenience is not sufficient justification to maintain a sex-based classification; neither is habit and comfort with a traditional classification).


\textsuperscript{228} Adams, 919 F. Supp. at 1504 (addressing safety concerns and finding them insufficient to justify exclusion of girl from boys’ wrestling because the policy of exclusion is not substantially related to the safety objective); Leffel v. Wis. Interscholastic Athletic Ass’n, 444 F. Supp. 1117, 1122 (E.D. Wis. 1978) (determining that the safety concern is a valid objective, but finding that the exclusion of girls from boys’ team is not substantially related to the objective, thus violating the Equal Protection Clause).

\textsuperscript{229} For example, Babe Didrickson was recently named one of the best athletes in the world of all time. See Bob Stratton, A Babe Shall Lead Them; Pioneers of the LPGA Recall the Tour’s Birth Fifty Years Ago When Babe Didrikson Zaharias Won the First U.S. Open, HAMILTON SPECTATOR,
the concern is for weaker or smaller students who may get hurt, then the restriction should be based on strength or size of the athlete. This would restrict weak and vulnerable males, as well as perhaps average females, from playing with stronger or bigger players, while allowing students of both sexes who are capable of playing with low injury risk to play.230 If the true concern is for student safety, there are other ways to address those concerns, such as improving training, changing rules, improving equipment, and having separate teams for players with different strength, skill, body weight, and/or maturity.231

2. Concerns that women will get squeezed out of sports do not meet strict scrutiny

Another concern is that women will get pushed out of playing sports if all sports teams became co-ed.232 Because, on average, men are stronger and bigger than women, the fear is that men will displace women if they are allowed to compete on the same teams as women.233 However, there are sex-neutral ways of ensuring that women will have equal opportunities to play sports as men, and these sex-neutral ways should be used instead.234 Title IX, by requiring that proportional numbers of men and women are engaged in sports at a school, and that a proportional amount of money is spent towards athletes of both sexes, is an excellent way to ensure that women will have equal opportunities as men to play sports.235 For example, a school with many athletic boys and not many athletic girls could start programs for beginner athletes, and encourage all students to get involved: this action, while not discriminating against either boys

July 13, 2000 at E6. See also Robinson, supra note 40, at 352-53 (listing women who have played sports with and at the same level as some of the best men). There is significant overlap in size and sports ability between men and women. See supra note 68.

230. Adams, 919 F. Supp. at 1504 (explaining that boys as well as girls are in danger of injuries due to playing sport, and it is improper to subject boys to greater risk of injury than girls); Force, 570 F. Supp. at 1028-29 (excluding a girl from a football team is not well tailored to the governmental interest of safety for students when males prone to injury are allowed to try out for football team but all girls are not); Hoover v. Meiklejohn, 430 F. Supp. 164, 169 (D. Colo. 1977) (“The failure to establish any physical criteria to protect small or weak males from the injurious effects of competition with larger or stronger males destroys the credibility of the reasoning urged in support of the sex classification.”).

231. B. Glenn George, Fifty/Fifty: Ending Sex Segregation in School Sports, 63 OHIO ST. L.J. 1107, 1149 (2002) (describing how the dangers of football led to changes in rules to make it safer). See also Malina, supra note 220, at 204 (describing how strength and size disparities between large and small boys in contact force can increase chance of injury for the smaller boys).

232. Robinson, supra note 40, at 352 (explaining that if all single sex teams were converted into co-ed teams, less women would be able to play because women in general have a disadvantage in sports when compared to men).


235. See Robinson, supra note 40, at 321-22 (opining that Title IX and the Fourteenth Amendment could “prosper as complementary [protectors] . . . of women’s rights in sports”).
or girls, would have the effect of helping girls at the school achieve substantive equality in sports opportunity. Alternatively, the school could support a sport in which women generally have an advantage over men, or in which women generally have a higher interest in than men.

3. Redressing past discrimination against women does not meet strict scrutiny

Another justification for maintaining sex-segregated sports teams, and in particular, for not allowing boys to play on girls’ teams, is to redress past discrimination against women. Boys have had greater opportunity to participate in sports in the past, and still participate in greater numbers than girls. However, this justification is unfair when applied for the benefit of girls who have not been disadvantaged, and to the detriment of boys who have not been advantaged. Further, this disadvantages smaller, weaker, less skillful boys, who may be at a good level to play with the majority of the girls, but not with the majority of the boys.

Under the strict scrutiny analysis, the Supreme Court has found that laws and policies cannot advantage one race above another on the justification that one race faced disadvantage in the past. Therefore, under the strict scrutiny

---

236. See supra note 8 and accompanying text.
237. See Attorney Gen. v. Mass. Interscholastic Athletic Ass’n, 393 N.E.2d 284, 293 n.34 (Mass. 1979) (stating that women may have an advantage over men in sports that test balance because of their lower center of gravity). See generally Jennifer Hargreaves, Sporting Females: Critical Issues in the History and Sociology of Women’s Sports 285-86 (Routledge, 1994) (stating that women generally have advantages over men in sports which emphasize endurance and flexibility, whereas men have an advantage over women in sports which emphasize speed and strength).
238. Such information could be gathered, for example, by survey.
239. See, e.g., Kelley v. Bd. of Trs., 35 F.3d 265, 272 (7th Cir. 1994) (holding that Title IX does not violate equal protection guarantees because Congress has broad powers to make up for past discrimination); Clark v. Ariz. Interscholastic Ass’n, 695 F.2d 1126, 1131 (9th Cir. 1982) (upholding a rule that did not allow a boy to play on a girls’ volleyball team on the basis that it redressed past discrimination against girls); Kleczek v. R. I. Interscholastic League, Inc., 768 F. Supp. 951, 956 (D.R.I. 1991) (recognizing that redressing the historically disparate opportunities in sports for men and women is a significant governmental objective).
240. See Woods, supra note 233, at 896.
241. Indeed, the analogous justification is not permitted in the race context. See infra note 243 and accompanying text.
242. See Mass. Interscholastic Athletic Ass’n, Inc., 393 N.E.2d at 290 (“Disadvantages suffered by males are often premised on a ‘romantic paternalism’ stigmatizing to women.”).
analysis, stating that women and girls have been disadvantaged in the past is not a sufficient justification for upholding segregated sports teams.244

Under a strict scrutiny regime, public schools and other government buildings would change in an additional way: bathrooms would no longer be sex-segregated.

C. Bathrooms would be integrated

Under a strict scrutiny test for sex, most if not all, state-sanctioned sex segregation would end, including maintaining separate bathrooms for men and women. These legalized differences only serve to reinforce existing sex stereotypes and ideas of difference.245 On a daily basis people must clearly show which sex they are, so that others know how to classify them, in speech and in knowing which bathroom the person belongs in. On a daily basis, people, including transgender and transsexual people, who do not fit into these clear gender boxes are made to feel like they do not fit in, and are pressured to conform to sex stereotypes.246 Common concerns with sex-integrated bathrooms include concerns about comfort, privacy, and safety.247 None of these concerns would justify sex-segregated bathrooms under a strict scrutiny test.

1. Comfort, convenience, and privacy concerns do not meet strict scrutiny

While it is true that most males will have an easier time than most females using a urinal because of physical differences in anatomy,248 there is no reason why all bathrooms cannot have both urinals and stalls, to accommodate both men and women. However, there would be no need to renovate current bathrooms (a large monetary expense). In current women’s bathrooms that have no urinals, men can use the stalls. In current men’s bathrooms that have urinals, women can use the stalls, while men can continue to use the urinals. At first men and women may be uncomfortable by the fact that now women may catch a glimpse of a man’s penis on entering a bathroom if it has urinals. Men wanting especially to

---

245. See supra Part II:B.
246. See supra Parts II:B:3 & II:C:3.
avoid this could use a stall while women could make sure to avert their eyes. In time, people would get used to the situation.

For example, a few years ago I visited my sister at her college where she lived in a housing unit with co-ed bathrooms, including co-ed group shower facilities. As I was taking a shower, a naked man I had never met walked in and began showering under a shower-head a few feet away. While I was initially taken aback, I managed to complete my shower and get clean with no problems. Habit and discomfort with change are not sufficient justifications under the strict scrutiny standard, because they are not “compelling.”  

2. Safety concerns do not meet strict scrutiny

Safety reasons have long been cited to maintain discrimination and keep men and women separate. However, because there is significant overlap in the vulnerability levels of men and women (not all women are more vulnerable than all men), this justification cannot meet strict scrutiny because it is not narrowly tailored. If a bathroom is a dangerous place, then people of the same sex as the potential attacker are also in danger, and the government or institution should take steps to alleviate the dangerous conditions for the potential same-sex as well as opposite-sex victims. If there are higher incidents of harassment because of co-ed bathrooms, as some posit, then safety measures such as surveillance cameras, outside the stalls or focusing on the entrance-way, could be installed. However, the inconvenience of entering a bathroom marked “women” will not

---

249. In fact, it is likely these justifications would not pass the intermediate scrutiny test. See Craig v. Boren, 429 U.S. 190, 198 (1976) (recognizing that habit and comfort cannot be used to maintain a stereotypical classification).

250. See, e.g., Muller v. Oregon, 208 U.S. 412, 421 (1908) (upholding a law which did not allow women to work as long hours as men because of, inter alia, safety concerns); Bradwell v. Illinois, 83 U.S. 130, 141 (1872) (Bradley, J., concurring) (citing females’ customary need for male “protection” as a reason not to admit them to the bar).

251. See supra note 68. Cf. Adams v. Baker, 919 F. Supp. 1496, 1504 (D. Kan. 1996) (explaining that boys as well as girls are in danger of injuries due to playing sport, and it is improper to subject boys to greater risk of injury than girls); Force v. Pierce City R-VI School Dist., 570 F. Supp. 1020, 1029-30 (W.D. Mo. 1983) (excluding a girl from a football team is not well tailored to the interest of student safety when males prone to injury may try out for a football team but all girls may not); Patricia B. Campbell & Jo Sanders, Challenging the System: Assumptions and Data Behind the Push for Single-Sex Schooling, in GENDER IN POLICY AND PRACTICE: PERSPECTIVES ON SINGLE-SEX AND COEDUCATIONAL SCHOOLING 31, 40 (Amanda Datnow & Lea Hubbard, eds., RoutledgeFalmer, 2002) (explaining that the real problems in schools are classroom misbehavior and disrespect, removing girls does nothing to protect boys who are also harassed by that behavior).

252. See Brett M. Kyker, “Initial Consent” Rape: Inherent and Statutory Problems, 53 CLEV. ST. L. REV. 161, 162 n.11 (2006) (showing that between 5%-10% of rape victims are male rape victims of other males); cf. Hoover v. Meiklejohn, 430 F. Supp. 164, 169 (D. Colo. 1977) (“The failure to establish any physical criteria to protect small or weak males from . . . [injury] destroys the credibility of the reasoning urged in support of the sex classification.”).

253. That way people’s privacy inside the stalls and by the urinals could be maintained while capturing who was inside a bathroom at the time of an incident.
necessarily deter a person intent on entering a bathroom to harass, assault, or rape. Instead, potential male offenders may well be deterred from attacking inside a bathroom that they know another male could enter at any time. Indeed, there are already co-ed bathrooms across the country in college dorms with no reports of decreased safety or increased violence.254

Sex-specific language use is another area where a seemingly innocent classification has harmful effects255 and would have to change under a strict scrutiny standard.

D. Language would change by the use of sex-neutral pronouns

The English language illustrates the pervasiveness of sex classifications in our society and perpetuates and normalizes stereotypes and conceptions of difference.256 English language pronouns are sex-specific: a person wanting to use a singular pronoun for a person must either use “he” or “she.” This unconsciously makes it frequent and normal to classify people by sex, even if a person’s sex has nothing to do with the point of the conversation, reference, or thought.257

If sex-based classifications were judged with strict scrutiny, government actors would not be permitted to classify people based on their sex, because there is no compelling reason to justify such sex classification.258 One consequence would be that public school teachers would not classify students by
using the pronouns “he” and “she” to talk with children in the classroom, unless sex was relevant to the conversation. Sex-neutral pronouns such as "se" (pronounced “see”) would be used instead of “he” or “she,” and "hir" (pronounced “here”) would be used instead of “his” or “her”.259 To see why, consider the following hypothetical. If Johnny makes a comment in class about poetry, a teacher would not ask another student, “What did you think of black Johnny’s comment?” This would be labeling Johnny with the label “black,” even though race is irrelevant to the conversation. By contrast, it would not be unusual for the same teacher to ask another student, “What did you think of his comment?” By using the word “his,” the teacher is labeling Johnny with the label “male,” even though sex is irrelevant to the conversation. There is no compelling justification to uphold the use of this type of sex-based classification in language. Habit is likely the only justification, and is not sufficient.260

If sex stereotypes are to be broken down, there should be no need to classify someone by sex every time a reference to that person is made. Although this is contrary to how people have learned to speak the English language, after some time it would become normal, similar to how such gendered phrases as “workman’s compensation” and “chairman” have changed to the sex-neutral phrases of “worker’s compensation” and “chairperson.” By changing pronoun use, sex stereotypes and subconscious sex classifications will be weakened.

IV. UNDER THE PROPOSED REGIME, FEW SEX CLASSIFICATIONS WOULD REMAIN

De facto segregation and rare instances of government sex segregation would be permitted even under a strict scrutiny regime.

A. De facto segregation may still exist

De facto segregation may exist for some time after legal segregation is abolished. More boys than girls may choose to join the football team, and more girls than boys may choose to join the cheerleading squad. However, an individual boy or girl will have more options to act outside the norm. Similarly,

259. See Feinberg, supra note 86, at 80 (advocating sex-neutral pronouns as a way to counter America’s need to know the answer to “is it a boy or a girl?” and to thereby enable individuals the freedom to discover their own identity, beyond the confines of sex or gender); see also John M. Ohle, Constructing the Trannie: Transgender People and the Law Footnote, 8 J. GENDER RACE & JUST. 237, 239 (2004) (advocating the use of sex neutral pronouns “ze” and “hir” in society, including legal writing, to “free us all from . . . narrow understandings of sex and gender . . .”). Government speech would include such things as language of public teachers, government documents such as tax forms, etc. Additionally, people would not be required to provide information on their newborn’s sex at birth, or at any time.

most people will probably still choose to enter opposite-sex marriages than same-sex marriages. Given the choice, men might favor the bathrooms that have more urinals in them. Finally, many people may still choose to identify with sex-specific pronouns in their private lives. Indeed, even though racial segregation in schools is officially abolished, de facto segregation still occurs in many lunch rooms, due to residual social factors and individual preferences in deciding who to associate with and what groups to join. Similarly, sex-based classifications will still occur in society due to individual preferences; however, the crucial difference is that the government will not be mandating these differences.

B. Rare situations may justify sex-based classifications

In rare situations, a sex-based classification may be constitutional, if it is narrowly tailored to further a compelling governmental aim and there is no sex-neutral way to accomplish the same aim. However, this would likely only arise in cases of remedying specific acts of past discrimination against an individual, or to protect the life of a human being in an emergency situation. To ensure that sex-based classifications are noticed when they occur, a court should analyze whether the challenged classification would be permitted if allowed for race. If the answer is no, then the court should critically examine whether the classification needs to be made for sex. Only if there is a compelling government interest at stake, and the classification is narrowly tailored to meet that interest should the classification be upheld.

CONCLUSION

Any sex-based classification should be ruled unconstitutional unless it can meet the strict scrutiny test. Strict scrutiny applied to race requires that a race-neutral statute be used whenever possible and even “narrowly drawn racial classifications [may not be used] except as a last resort.”

261. See Sara Rimer, Learning to Live Peacefully in High School Melting Pot, N.Y. TIMES, May 17, 1984, at B1 (describing de facto racial segregation that occurs in many school cafeterias); cf. Carolyn B. Ramsey, Subtracting Sexism from the Classroom: Law and Policy in the Debate Over All-Female Math and Science Classes in Public Schools, 8 TEX. J. WOMEN & L. 1, 43-44 (1998) (stating that de facto segregation of girls from boys often occurs in school cafeterias and in lunch time games).

262. See supra note 28 and accompanying text.

263. In many cases sex-based classifications are not noticed by courts. See, e.g., Andersen v. King County, 138 P.3d 963, 988 (Wash. 2006) (en banc) (holding that the opposite-sex marriage requirement does not classify based on sex; opposite-sex marriage requirements treat men and women identically because neither may marry a person of the same-sex); see also supra notes 173-176 and accompanying text.

rights and responsibilities based on immutable physical characteristics rather than individual merits, interests, or abilities. Furthermore, sex-based classifications are incidents of our cultural history of prejudice and discrimination based on sex, just as race-based classifications are incidents of our cultural history of prejudice and discrimination based on race.

Adopting the strict scrutiny test for sex discrimination would be a big step forward to achieving sex equality and creating a society where people are able to develop their individual capacities and abilities free of the pressure and legal burdens that come from stereotypes and prejudice.
COUPLING RESPONSIBILITY WITH LIABILITY: WHY INSTITUTIONAL REVIEW BOARD LIABILITY IS GOOD PUBLIC POLICY

Danielle C. Beasley*

I. INTRODUCTION

The United States biomedical research system sets forth institutional review boards (IRBs) as an oversight structure for human subject research.1 In 1998, the Office of the Inspector General (OIG) of the Department of Health and Human Services issued a report, Institutional Review Boards: A Time for Reform, based on its one-year study of IRBs.2 In that report, the OIG sharply criticized IRBs for being overworked, lacking sufficient expertise for adequate review, having a bias in favor of research, and providing inadequate training for their members.3 Since that time, IRBs have faced an enormous amount of public pressure.

In 2002, the pressure on IRBs elevated to unprecedented levels because two cases, for the first time in history, named individual IRB members as defendants in a lawsuit. In Robertson v. McGee, twelve IRB members were named individually as defendants.4 The case “involved a melanoma cancer research study conducted by defendant Dr. Michael McGee at the University of Oklahoma Health Science Center-Tulsa campus.”5 The plaintiffs all “suffered from melanoma cancer and agreed to participate in the study, which was the subject of the litigation.”6 The plaintiffs contended that subject matter jurisdiction was proper based on their federal constitutional right to be treated

---

* Ms. Beasley is an attorney at Vercruysse Murray & Calzone, P.C. in Bingham Farms, Michigan. She received her J.D. from the University of Maryland School of Law and her B.A., with distinction, from the University of Michigan. The opinions expressed in this article are solely the author’s.

1. See Grimes v. Kennedy Krieger Inst., Inc., 782 A.2d 807, 813 (Md. 2001) (noting that an IRB is required “by either federal or state regulation, or sometimes by the conditions attached to the governmental grants that are used to fund research projects”).


5. Id.

6. Id.
with dignity, as well as from the federal regulations on the protection of human subject research. The issue of IRB member liability was not decided by the court in Robertson because the case was dismissed for lack of subject matter jurisdiction. Similarly, in Townsend v. University Hospital-University of Colorado, the individual members of the IRB, from January 1, 1998, to May 1999, were named as defendants. In Townsend, the plaintiffs filed suit in Texas state court listing multiple causes of action, including lack of informed consent, in relation to the death of Julia Caren Townsend Olivares while in the defendants’ care in Colorado. Like in Robertson, the issue of IRB member liability was not decided; the appellate court affirmed the dismissal of the case for lack of personal jurisdiction.11 Even though these cases were dismissed, the issue of IRB liability persists because other lawsuits listing an IRB or its members as defendants have been subsequently filed.12

Many commentators assert that IRB liability is problematic as it will result in difficulty recruiting IRB members, increased research costs, and defensive tactics by IRB members out of fear of potential litigation against them. This article counters these claims and argues that subjecting IRB members to tort liability is good public policy.

Section II of this article addresses the historical perspective of self-regulated research in the United States, including its highlights and concerns; the United States government’s regulatory oversight of human subject research; and recent biomedical research tragedies which have resulted in severe criticisms of IRBs. Section III details the fundamentals of an IRB negligence action. Section IV argues that holding IRBs liable for negligence that results from their shortcomings will serve as an efficient catalyst to improve the protections for human subject research participants. Section V concludes that our civil tort
system, which works in the context of rights and responsibilities, is a much needed vehicle to combat the problems of professional self-regulation and governmental oversight inadequacies.

II. HISTORICAL OVERVIEW

The publication of the Nuremberg Code in 1947 had little effect on the conduct of human experimentation in the United States in the years after World War II. Although the Nuremberg Code was certainly known to United States investigators, they paid minimal attention to it and continued carrying out harmful self-regulated research on subjects incapable of giving consent. The Tuskegee Syphilis Study, the Willowbrook Study, and the Jewish Chronic Disease Hospital Study aptly illustrated the problems of this self-regulated research. In response to these studies, Congress passed the National Research Act of 1974. However, even after formal regulations on human subject research were enacted in the United States, the long history of abuses has continued. As a result, the public’s attention has shifted to IRBs.

A. Self-Regulated Research in the United States

Spanning four decades, the Tuskegee Syphilis Study is perhaps the most infamous self-regulated biomedical research abuse in the United States. In 1932, the United States Public Health Service (USPHS) initiated the Tuskegee Syphilis Study to document the effects of untreated syphilis. The subjects of the investigation were approximately four hundred poor, African American men from Macon County, Alabama. “When penicillin became widely available by the early 1950s as the preferred method of treatment for syphilis, the men did not receive” the medicine. “In fact, on several occasions, the USPHS actually

14. On December 9, 1946, an American tribunal began its criminal proceedings against Nazi doctors for experiments carried on during World War II using unconsenting concentration camp victims as research participants and subjecting them to great pain and suffering, disfigurement, and death. See JONATHAN D. MORENO, UNDUE RISK: SECRET STATE EXPERIMENTS ON HUMANS 57 (W.H. Freeman and Company 1999) (2000). “The final section of their judgment includes ten points that quickly became known as the Nuremberg Code, which begins, ‘The voluntary consent of the human subject is absolutely essential,’ and proceeds to establish a number of other [ethical] requirements.” Id. at 80.

15. Throughout this article, the term “self-regulation” is used to describe a context where people or institutions operate without accountability for their actions because governmental oversight is virtually non-existent. Given this essentially “lawless” context, these people or entities are left to regulate themselves by relying on their own judgment to act appropriately.


17. See discussion infra Section II.C.


19. Id.

20. Id.
sought to prevent [the penicillin] treatment."\textsuperscript{21} In 1972, accounts of the study appeared in the national press causing the Department of Health, Education and Welfare to close down the experiment.\textsuperscript{22}

Another frequently cited self-regulated study took place between 1956 and 1971 at the Willowbrook State School, “a state facility that housed profoundly impaired children and adolescents.”\textsuperscript{23} “Trying to find a way to protect people from hepatitis, [the investigators] deliberately infected some of the children with the virus.”\textsuperscript{24} In an attempt to defend the study, investigators claimed to have consent from parents of all subjects involved in the research.\textsuperscript{25} However, many have claimed that the parents were coerced into agreeing so that their children would gain admission to the facility.\textsuperscript{26}

The 1963 Jewish Chronic Disease Hospital Study is a third illustrative self-regulated case on United States biomedical research abuses. In this study, the investigator injected live cancer cells into chronically ill, unconsenting patients.\textsuperscript{27} The investigator defended the study by “noting that no patient was at risk of developing cancer because the[ir] immune system would ultimately reject the foreign cells” even though they were ill.\textsuperscript{28}

Eventually, the United States federal government responded to the publicity concerning these self-regulated research abuses by establishing oversight regulations on human subject research.

\textbf{B. The Oversight Structure for Human Subject Research is Created}

The numerous biomedical research abuses in the United States led to the passage of the National Research Act of 1974, which called for the establishment of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research.\textsuperscript{29} In 1979, the Commission published recommendations known as the Belmont Report.\textsuperscript{30}

[\text{The Belmont Report articulated three [ethical] principles [for human research conduct]: respect for persons (the recognition of the right of persons to exercise autonomy), beneficence (the minimization of risk incurred by research subjects and the maximization of benefits to them}]

\begin{itemize}
\item 21. \textit{Id}.
\item 22. \textit{Id}.
\item 23. MORENO, supra note 14, at 249.
\item 24. \textit{Id.} at 250.
\item 25. \textit{Id}.
\item 26. \textit{Id}.
\item 27. \textit{Id.} at 246.
\item 28. \textit{Id}.
\item 30. \textit{See} Harold Varmus & David Satcher, Ethical Complexities of Conducting Research in Developing Countries, in \textit{Tuskegee Truths: Rethinking the Tuskegee Syphilis Study} 584 (Susan M. Reverby ed., University of North Carolina Press 2000).
and others), and justice (the principle that therapeutic investigations should not unduly involve persons from groups unlikely to benefit from subsequent applications of the research).

Presently, those principles serve as the foundation for the federal regulations governing human subject research. Additionally, the National Research Act of 1974 mandated that human subject research entities provide assurances to the Secretary of Health, Education and Welfare that they have established an institutional review board in order to protect the rights of human subjects involved in the research.

“Until 1991, federal departments and agencies that conduct, support, or regulate research used a variety of policies and procedures to protect human research subjects.”33 “To eliminate confusion and to promote uniformity,” sixteen agencies have adopted the Federal Policy for the Protection of Human Subjects, which is referred to as the Common Rule.34 There are three components to the Common Rule: assurances, institutional review boards, and informed consent.35 First, the Common Rule mandates that all federally funded human subject research institutions provide a written assurance to the sponsoring agency that it will comply with all requirements enumerated in the regulations.36

Next, the Common Rule requires that all federally funded research institutions establish an IRB with at least five members, including men and women, with different backgrounds.37 The IRB’s main functions are to review initial and continuing research.38 Initial review entails consideration of risks and benefits, equitable selection of subjects, adequate informed consent, data and safety monitoring, protection of subjects’ privacy, and additional safeguards for specific vulnerable populations.39 The IRB can approve, disapprove, or require modifications to research.40 Once research has passed the muster of initial review, the regulations require that the research be reviewed by the IRB at least once a year or at “intervals appropriate to the degree of risk.”41

31. Id. at 585.
34. Id.; see also Resnik, supra note 13, at 137.
35. See Stalup, supra note 3, at 1599.
37. Id. § 46.107; see also Resnik, supra note 13, at 137 (“IRBs should be composed of members who have competence to review biomedical and behavioral research (scientific members), as well as members who have knowledge of applicable laws, institutional policies, and professional standards of conduct (non-scientific members.”).
38. 45 C.F.R. § 46.109(a); see also Resnik, supra note 13, at 137.
39. See generally 45 C.F.R. § 46.111; see also Resnik, supra note 13, at 137.
40. 45 C.F.R. § 46.109(a).
41. Id. § 46.109(e); see also Resnik, supra note 13, at 137-38 (asserting the more risky the research, the greater level of monitoring it will require).
Finally, the Common Rule necessitates adequate informed consent before human beings can serve as research subjects. As part of the informed consent process, prospective participants must be given an explanation of any foreseeable risks, benefits, and appropriate alternatives; a statement pertaining to research confidentiality; a list of contacts for research-related questions; and a statement on the voluntariness of research and permissive subject withdrawal. Furthermore, no informed consent form can include language that waives or appears to waive any of the participant’s legal rights or that appears to release the investigator, sponsor, institution, or its agents from liability.

However, even after the enactment of the Common Rule, biomedical research abuses continued to occur in the United States. As a result, in recent years, some IRBs have been criticized for deficiencies in their oversight of certain research projects.

C. Research Tragedies Resulting in Scrutiny of IRBs

The public scrutiny on IRBs dramatically intensified on September 17, 1999, when Jesse Gelsinger died in a Phase I gene therapy trial at the University of Pennsylvania’s Human Gene Therapy Institute. Gelsinger suffered from a mild form of a rare metabolic disorder called ornithine transcarbamylase deficiency (OTC). As part of the study, genetic material was injected into Gelsinger’s bloodstream, which triggered multiple organ system failures resulting in his death. After Gelsinger’s death, the Office for Human Research Protections (OHRP) criticized the University of Pennsylvania IRB by identifying

42. 45 C.F.R. § 46.116.
43. See id. The informed consent forms must be written or verbalized in a manner where the subject or his/her representative can understand.
44. See id.
45. Id.
46. Phase I studies involve the initial introduction of an investigational new drug into humans. 21 C.F.R. § 312.21(a) (2006) (“These studies are designed to determine the metabolism and pharmacologic actions of the drug in humans, the side effects associated with increasing doses, and, if possible, to gain early evidence of effectiveness.”).
47. Resnik, supra note 13, at 132.
48. See JOSEPH PAMNO, GENE THERAPY: TREATING DISEASE BY REPAIRING GENES 46 (Facts on File, Inc. 2005).
49. See id. at 51-52.
numerous deficiencies regarding its approval of the gene therapy trial. The IRB’s deficiencies included the failure to obtain sufficient information when reviewing a protocol revision and approval of poor informed consent documents.

In July 2000, the University of Oklahoma Health Sciences Center’s federally funded research was suspended because it failed to give adequate protections for human research subjects during a research study testing a vaccine for melanoma cancer. The OHRP cited a number of findings for the study stating, *inter alia*, that the IRB failed to ensure that additional safeguards were used on the vulnerable population of terminally ill subjects, did not satisfy the requirements for continuing review, and should not have approved inadequate informed consent documents. Eventually, this research study became the subject of litigation in Oklahoma and all of the IRB members from the study were individually named as defendants in the lawsuit. The plaintiffs asserted that the defendants failed to comply with federal regulations and to notify the plaintiffs of this failure, thereby amounting to “involvement in the study . . . without their consent.” The court dismissed the suit for lack of subject matter jurisdiction.

In 2001, John Hopkins University became involved in its own research abuses. Ellen Roche was a healthy, twenty-four year old who volunteered to participate in an asthma study at John Hopkins University. As a participant in the study, Roche inhaled hexamethonium, a substance that had not been approved by the FDA for use in humans and had never been approved by the


52. See id.

53. See Letter from M.A. Carome, Chief Compliance Oversight Branch, Div. of Human Subject Prots. to J.J. Ferritti, Senior Vice President and Provost, Univ. of Okla. Health Scis. Ctr., and N.L. Nisbett, Dir. of Research Admin., Univ. of Okla. Health Scis. Ctr. (July 7, 2000), http://www.hhs.gov/ohrp/detrm_letrs/jul00a.pdf (regarding Human Research Subject Protections Under Multiple Project Assurance (MPA) M-1448 and stating that “OHRP acknowledges [UOHSC’s] report that 26 federally supported projects being conducted by the UOHSC-Tulsa have been suspended, as well as all other research involving human subjects”).


56. Id. at *2.

57. The court found that no private right of action existed for alleged violations of international laws, such as the Declaration of Helsinki and the Nuremberg Code. Id. at *3.

FDA for administration by inhalation. Shortly after beginning the study, Ellen Roche died. After her death, the OHRP cited the John Hopkins IRB for significant violations of federal regulations for inadequately reviewing the experiment, neglecting to acquire literature “available via routine MEDLINE and Internet database searches” on the known lung toxicity of hexamethonium, and approving an informed consent document that failed to adequately describe the research procedures and reasonably foreseeable risks and discomforts.

Also in 2001, a John Hopkins-affiliated research institution, the Kennedy Krieger Institute, was a defendant in a landmark decision by Maryland’s highest state court. Grimes v. Kennedy Krieger Institute, Inc. involved a research study to determine how effective varying degrees of lead paint abatement procedures were. The research institute encouraged landlords to rent apartments with lead dust to low-income families with children. The researchers periodically tested the children’s blood to compare the lead contamination in their bloodstream with the contamination levels of lead dust in the homes. The IRB approved the informed consent document even though it did not contain a complete and clear explanation that the research was designed to evaluate the success of the abatement procedures by measuring the extent to which the children’s blood was being contaminated. The Maryland Court of Appeals found for the plaintiffs. Although the IRB was not named as a defendant, the court criticized the IRB stating that it had “abdicated . . . responsibility, instead suggesting to the researchers a way to miscast the characteristics of the study in order to avoid the responsibility inherent in nontherapeutic research involving children.”

This overview demonstrates that private individuals, government agencies, and courts have criticized IRBs for failing to heed their role to safeguard the welfare of research participants and ensure that human research studies comply with federal regulations. In response, some have argued that the IRB system needs to be dramatically changed to meet the demands of the current research environment. Others have asserted that increased regulatory oversight of IRBs,
including an accrediting system, is the solution. This article emphasizes that coupling responsibility with liability is the best method to protect human research participants.

III. IRB NEGLIGENCE ACTIONS

“Traditionally, IRBs have rarely been named in lawsuits.” Recently, however, this landscape has changed as IRBs have entered the litigation spotlight in research-related lawsuits. Plaintiffs who desire to file suit naming the IRB as a defendant can either name the entire board collectively as a defendant or name individual IRB members as defendants. In either regard, a plaintiff is most likely to file a negligence suit against the IRB or its members. Consequently, this article will exclusively address that claim.

A. Elements of a Negligence Claim

The plaintiff has the burden of proving all the elements by a preponderance of the evidence in negligence cases. To maintain a negligence action, the plaintiff must assert the following elements: (1) the defendant owed a duty to the plaintiff, (2) the defendant breached the duty, (3) the defendant’s breach caused the plaintiff’s injuries, and (4) the plaintiff suffered damages.

1. Duty of Care

The threshold question in an IRB negligence case is whether the IRB has a legal duty of care towards research subjects. A duty can be “established by a legislative enactment or administrative regulation which so provides,” and the federal regulations appear to create the standard of conduct for IRBs toward human research subjects. At least two courts have yielded opinions which support this conclusion.

In Gregg v. Kane, plaintiffs sued Dr. Daniel M. Kane, Dr. Stephen L. Trokel, Visx Inc., and the Wills Ever Hospital for negligence and lack of informed

69. Id.; see also discussion infra Section IV.A.2.
70. Hoffman & Berg, supra note 13, at 381.
71. See id. at 381-82.
75. RESTATEMENT (SECOND) OF TORTS § 285 (1965); see also Steed v. Grand Teton Council of the Boy Scouts of Am., 172 P.3d 1123, 1128 (Idaho 2007) (“Statutes and administrative regulations may define the applicable standard of care owed.”).
76. See Anderlik & Elster, supra note 72, at 224 (stating that in the United States, “the federal regulations arguably set forth the standard of conduct to which IRBs should be held”); see also Resnik, supra note 13, at 143 (“Federal regulations appear to constitute a source for legal duties for IRB members.”).
consent relating to laser eye surgery.\textsuperscript{77} The defendants filed seven motions for summary judgment, but all were dismissed by the court.\textsuperscript{78} In its unpublished order, the federal court stated that the Wills Ever Hospital IRB had duties under the federal regulations because “the FDA regulations make IRBs . . . responsible for insuring that informed consent will be sought from all prospective subjects, in accordance with fairly detailed standards governing such consent.”\textsuperscript{79} The court went on to assert that because the consent for one of the plaintiffs was defective, “Wills could be found liable for not fulfilling its FDA-mandated responsibilities.”\textsuperscript{80} Given that Wills could be found liable, it follows that the IRB could also be found liable because the FDA regulations governed the conduct of the IRB.\textsuperscript{81}

Additionally, in \textit{Grimes}, the Maryland Court of Appeals held that “informed consent agreements in nontherapeutic research projects, under certain circumstances can constitute contracts; and that, under certain circumstances, such research agreements can, as a matter of law, constitute ‘special relationships’ giving rise to duties, out of the breach of which negligence actions may arise.”\textsuperscript{82} The court went on to hold that “governmental regulations can create duties on the part of researchers towards human subjects.”\textsuperscript{83} Similarly to \textit{Gregg}, the court’s reasoning supports the possibility that the federal regulations also create a duty on IRBs towards research participants. Therefore, in an IRB negligence action, a court could likely find that the IRB or IRB member owes the plaintiff a duty of care.

2. Breach of the Standard of Care

In a negligence action, if the defendant’s actions fall below the standard of care, the defendant will be deemed to have breached the duty he or she owed the plaintiff. For an IRB negligence case, a court will likely apply the negligence per se standard, the reasonable person standard, or the professional negligence standard.\textsuperscript{84}

Assuming \textit{arguendo} that the federal regulations govern IRB responsibilities, a negligence per se standard would be an appropriate standard to apply to an IRB negligence action.\textsuperscript{85} The theory behind negligence per se is that the violation of

\textsuperscript{78.} Id.
\textsuperscript{79.} Id. at *4.
\textsuperscript{80.} Id.
\textsuperscript{81.} Resnik, supra note 13, at 142.
\textsuperscript{83.} Id.
\textsuperscript{84.} See Resnik, supra note 13, at 148.
\textsuperscript{85.} See Grable & Sons Metal Prods., Inc. v. Darue Eng’g & Mfg., 545 U.S. 308, 318 (2005) (“The violation of federal statutes and regulations is commonly given negligence per se effect in state tort proceedings.”). But see id. n. 6 (“Other jurisdictions treat a violation of a federal statute as evidence of negligence or . . . as creating a rebuttable presumption of negligence.”).
a statute or regulation is more than just evidence of negligence, but it is negligence itself. Accordingly, to prove negligence per se, a plaintiff needs to show first, that the defendant violated the regulation and, second, that the plaintiff was a member of the class of people that the regulation was designed to protect. Applying these principles to the human research context, a plaintiff might argue that the federal regulations were adopted to protect human beings from research-related abuses thereby making human research subjects the protected class. Thus, the federal regulations would govern the standard of care, and IRB conduct that falls below those guidelines would amount to negligence per se.

A court might also apply the reasonable person standard to an IRB negligence claim. The reasonable person standard is a hypothetical construct setting forth an objective standard based on what a prudent community member would have done under similar circumstances. Under this standard, a plaintiff might argue that a reasonable IRB member would understand and comply with the requirements stated in the federal regulations; if, however, the IRB member acts unreasonably in his or her assessments or otherwise disregards the regulations, then the IRB member would breach the duty of care owed to research subjects.

Furthermore, the court could apply the professional negligence standard to an IRB negligence lawsuit if it finds that participation on an IRB is itself a specialty. In professional negligence cases, the court looks at custom to determine the standard of care, and if the defendant’s actions do not correspond with the profession’s custom, then the defendant will have breached his or her duty. Thus, under the professional negligence standard, if an individual

86. See Martin v. Herzog, 126 N.E. 814, 815 (N.Y. 1920) (“The unexcused omission of the statutory signals is more than some evidence of negligence. It is negligence in itself.”) (emphasis omitted); see also Hendrix v. Schulte, 736 N.W.2d 845, 849 (S.D. 2007) (“The statute or ordinance becomes the standard of care or the rule of the ordinarily careful and prudent person.”).
87. HARRY SHULMAN ET. AL, LAW OF TORTS CASES AND MATERIALS 233 (Foundation Press 4th ed. 2003) (“The general rule for negligence per se limits its applications to cases in which the plaintiff is among the class of individuals that the statute was intended to protect and the injuries are of the character of injuries which the statute was designed to prevent.”); see also Hendrix, 736 N.W.2d at 849 (“The violation of a statute or ordinance, designed for the benefit of individuals, is of itself sufficient to prove such a breach of duty as will sustain an action for negligence brought by a person within the protected class if other elements of negligence concur.”).
88. Resnik, supra note 13, at 149.
89. “[C]ourts have gone to unusual pains to emphasize the abstract and hypothetical character of this mythical person.” Radtke v. Everett, 501 N.W.2d 155, 166 (Mich. 1993) (citing PROSSER & KEETON, TORTS § 32, at 173-75 (5th ed. 1984)). The reasonable person is “not to be identified with any ordinary individual who might occasionally do unreasonable things.” Id. Instead, the reasonable person is “a prudent and careful person who is always up to standard.” Id. Thus, the reasonable person is “a personification of a community ideal of reasonable behavior determined by the jury’s social judgment.” Id.
90. Bryant v. LaGrange Mem’l Hosp., 803 N.E.2d 76, 84 (Ill. App. Ct. 2003) (“A professional is expected to use the same degree of knowledge, skill and ability as an ordinarily careful professional would exercise in similar circumstances.”); see also Brune v. Belinkoff, 235 N.E.2d
member of an IRB was named as a defendant, the court could assess deviation from the standard of care by determining whether the IRB member acted in conformity with how an average IRB member in the applicable professional field might perform under the given circumstances.

Even if the plaintiff can demonstrate that the defendant breached the applicable standard of care, the elements of causation and damages must be shown in order for the plaintiff to prove negligence.

3. Causation

In an IRB negligence action, the plaintiff must be able to demonstrate that there is a causal link between the IRB’s breach of the standard of care and the alleged injury. The requisite causal link can be established by cause-in-fact, which means that the injury or harm would not have occurred “but for” the IRB’s negligent conduct. Alternatively, if the case involves multiple defendants, the court may use the substantial factor test as a substitute for cause-in-fact. The substantial factor test evaluates whether each defendant’s “contributing cause, standing alone, is a substantial factor in causing the alleged injury.” Thus, to prevail under the substantial factor test, the plaintiff needs to prove that each IRB member made an independent but substantial contribution to the plaintiff’s injury.

In addition to proving cause-in-fact or causation under the substantial factor test, the plaintiff must put forth evidence on the issue of proximate cause. The general rule on proximate cause is that the harm to the plaintiff must be a reasonably foreseeable consequence of the defendant’s conduct. To make the

793, 798 (Mass. 1968) (“A specialist should be held to the standard of care and skill of the average member of the profession practicing the specialty.”).


92. “When several factors combine to produce an injury, and where any one of them, operating alone, would have been sufficient to cause the harm, a plaintiff may establish factual causation [i.e., cause-in-fact] by showing that the defendant’s actions, more likely than not, were a ‘substantial factor’ in producing a plaintiff’s injuries.” GKC Mich. Theaters, Inc. v. Grand Mall, 564 N.W.2d 117, 121 (Mich. Ct. App. 1997). “[I]n the great majority of cases, [the substantial factor test] produces the same legal conclusion as the but-for test.” Mitchell v. Gonzales, 819 P.2d 872, 879 (Cal. 1991) (emphasis omitted). The substantial factor test is useful in the disposition of troublesome cases involving multiple causes. See id. at 878-79.

93. Tragarz v. Keene Corp., 980 F.2d 411, 424 (7th Cir. 1992); Collins v. Li, 933 A.2d 528, 549 (Md. Ct. Spec. App. 2007) (“The substantial factor test was devised to address situations where two independent causes concur to bring about an injury, and either cause, standing alone, would have wrought the identical harm.”).


95. Overseas Tankship (U.K) Ltd. v. Mort’s Dock & Eng’g Co. Ltd. (The Wagon Mound No. 1) 1 All E.R. 404 (1961) (holding that the essential factor in determining liability for the
necessary proximate cause showing in an IRB negligence case, a plaintiff must demonstrate that the IRB’s actions are not too remote to be considered causally linked to the plaintiff’s injury thereby keeping the chain of causation in tact.96

4. Damages

Provided that the plaintiff has established duty, breach, and causation by a preponderance of evidence in an IRB negligence action, the plaintiff can recover damages to compensate them for the loss or injury caused by the IRB’s conduct.97 The goal is to restore the plaintiff, as nearly as possible, to the position he or she would have been in had the wrongful conduct not occurred.98

B. Bases for an IRB Negligence Action

In research-related lawsuits against IRBs, negligence could be alleged on several bases. However, two of the most common negligence claims against IRBs will be for insufficiency of informed consent and for negligent approval of research design.99

1. Insufficient Informed Consent

The principle of informed consent was identified in the Nuremberg Code.100 Informed consent generally means that a potential research subject should have the ability “to exercise free power of choice, without the intervention of . . .

consequences of a tortious act of negligence is whether the damage is of such a kind as the reasonable man should have foreseen); see also Idbeis v. Wichita Surgical Specialists, P.A., 173 P.3d 642, 652 (Kan. 2007) (“Proximate cause is defined as that cause which in natural and continuous sequence, unbroken by an efficient intervening cause, produces the injury and without which the injury would not have occurred, the injury being the natural and probable consequence of the wrongful act.”) (citation omitted).
It is well settled that there can be no proximate cause where there has intervened between the act of the defendant and the injury to the plaintiff, an independent, intervening, act of someone other than the defendant, which was not foreseeable by defendant, was not triggered by defendant’s act, and which was sufficient of itself to cause the injury. See also Helmus v. Michigan Dep’t of Transp., 604 N.W.2d 793, 797 (Mich. Ct. App. 1999) (“Proximate cause is that which operates to produce particular consequences without the intervention of any independent, unforeseen cause, without which the injuries would not have occurred.”).
98. Id.
force, fraud, deceit, duress, over-reaching, or other ulterior form of constraint or coercion; and should have sufficient knowledge and comprehension of . . . the subject matter involved . . . to enable him to make an understanding and enlightened decision.”

The notion of informed consent has been adopted in the Common Rule and is a criteria that IRBs must evaluate before research on human subjects can commence. Because, under negligence per se, the federal regulations establish the standard of care for human subject research where research participants are the protected class, the IRB’s failure to obtain legally effective informed consent could result in a negligence action. The case of Grimes v. Kennedy Krieger Institute is illustrative of this point.

In Grimes, the plaintiffs sued because there was no complete nor clear explanation in the consent agreements that parents signed for their children stating that the research was designed to use the extent of contamination in the children’s blood as a proxy for the effectiveness of the lead abatement procedures. In this monumental decision, the court held that “informed consent agreements in nontherapeutic research projects . . . can constitute contracts; and that under some circumstances, such research agreements can, as a matter of law, constitute special relationships giving rise to duties of care,” which if breached by the researcher, can result in negligence actions. The court also disapproved of the IRB’s decision to approve the consent forms for the study. It follows from the court’s logic that a negligence suit could have also been brought against the IRB or its members for approving research that failed to meet the standards of informed consent in the Common Rule.

2. Negligent Approval of Research Design

An IRB that does not take the necessary steps to ensure that research protocols are legally, ethically, and scientifically fit could be liable for negligently approving a research design. The plaintiffs in Robertson illuminated this point when they filed suit naming all individual members of the IRB as defendants in the lawsuit. Alan Milstein, the plaintiffs’ attorney in the case, explained that the IRB members were named as defendants because “[f]ederal regulatory agents made specific reference to the inadequate job of supervision that the IRB did.”

101. Id.
103. 782 A.2d 807 (Md. 2001).
104. Id. at 813.
105. Id. at 858.
106. See id. at 813.
Also, shortly after Ellen Roche’s death from an asthma study at John Hopkins, the OHRP conducted an investigation that resulted in the identification of a number of weaknesses associated with the IRB’s approval of the research. Particularly, the OHRP found that the investigators failed to advise the IRB that hexamethonium was labeled “for laboratory use only” by the manufacturer and did not provide – *nor did the IRB request* – information about the pharmacology, toxicity, or safety in humans of inhaled hexamethonium. Therefore, the John Hopkins IRB failed to obtain adequate information with respect to evaluating the risks associated with the research protocol.

Given that federal regulations charge IRBs with duties to protect research participants from harm, inadequate or deficient review and approval of research is a breach of that duty. As such, if a plaintiff can put forth evidence on causation and damages, then an IRB or its members can be found negligent because it did not take the requisite precautions to ensure that the protocols it approved complied with the applicable regulations.

This section demonstrates that IRB tort liability is not a legal fiction as the IRB or its members can be found negligent in a court of law. In response to suit, IRBs members could always present evidence at trial to disprove the elements of negligence. Moreover, the IRB or its members could raise a number of affirmative defenses to eliminate or mitigate the amount of damages awarded to the plaintiff by asserting contributory negligence, comparative negligence, assumption of risk, or government immunity. Nonetheless, the exorbitant expenses of litigation are likely to preclude suits against IRBs. Furthermore, courts have rejected the argument that IRBs are not subject to the traditional doctrines of negligence in this context. See the discussion below of the doctrine of contributory negligence. Id.


110. Id.

111. See Resnik, supra note 13, at 148.

112. Gyerman v. U.S. Lines Co., 498 P.2d 1043, 1050-51 (Cal. 1972) (en banc) (citing *RESTATEMENT (SECOND) OF TORTS* § 463 (1965)) (defining contributory negligence as “conduct on the part of the plaintiff which falls below the standard to which he [or she] should conform for his [or her] own protection, and which is a legally contributing cause cooperating with the negligence of the defendant in bringing about the plaintiff’s harm”). See *Li v. Yellow Cab Co.*, 532 P.2d 1226, 1230 (Cal. 1975) (en banc) (“The plaintiff’s contributory negligence bars recovery against a defendant whose negligent conduct would otherwise make him liable to the plaintiff for the harm sustained by him.”).

113. “The draconian nature of the doctrine of contributory negligence as a total bar to plaintiff’s recovery—regardless of how slight the plaintiff’s total negligence—has long been recognized.” SHULMAN, supra note 87, at 426. “In recent decades, most states have replaced this doctrine with the concept of comparative fault, under which a plaintiff’s fault reduces, but often does not eliminate recovery.” Id.

114. Assumption of risk is used differently by courts, but in its simplest form means that the plaintiff has given his or her “consent to relieve the defendant of an obligation to exercise care for his or her protection, and agrees to take his or her chances as to injury from a known or possible risk.” *RESTATEMENT (SECOND) OF TORTS* § 496A cmt. c (1965). In some states, assumption of risk will bar recovery while in others it is absorbed into comparative negligence thereby only mitigating damages. See id. § 496A cmt. d.

115. IRB members of a state or federal agency may be able to argue that they should be granted some form of qualified government immunity. See Resnik, supra note 13, at 172. Qualified-immunity defenses “shield government agents from liability for civil damages insofar as their
costs of trial coupled with the fact that trials often injure professional reputations (even if the defendant is found not guilty) means that tort liability is a real possibility with potential consequences for IRB members. This possibility of holding IRB members liable has sparked a debate on whether liability for IRB members is good public policy.

IV. IRB LIABILITY ARGUMENTS

Some might argue that the idea of holding IRB members personally liable for their actions does not make good business sense. However, this article asserts that a policy which allows negligent IRB members to escape responsibility for their actions is even more troublesome.

A. Arguments Against IRB Liability

Most argue that subjecting IRB members to tort liability is not a desirable practice for a number of reasons. First, they assert that it will be harder to get people to serve on IRBs. Second, they argue that there are viable deterrence alternatives to the legal system. Third, they believe that IRBs will engage in defensive practices in response to potential liability. Finally, they declare that IRB liability will result in increased research costs.

1. Harder to Get People to Serve on IRBs

Critics of IRB liability assert that the fear of liability will dissuade people from serving on IRBs. These individuals argue that being on an IRB is in itself a stressful and time-consuming task, especially in light of the fact that most IRB members do not receive compensation for their work. Consequently, “[f]or some, the genuine threat of legal liability may be the last straw.” However, the individual meritorious IRB liability case creates the appropriate accountability context. IRB members undertake great responsibilities in fulfilling their duties under the Common Rule. “Our [United States] legal system couples responsibility with liability.” We expect that those who serve the members of the public be held to the principle not to injure by conduct does not violate clearly established statutory or constitutional rights of which a reasonable person would have known.” Behrens v. Pelletier, 516 U.S. 299, 305 (1996) (internal quotations omitted).

116. See Resnik, supra note 13, at 178.
117. E.g., Stalcup, supra note 3, at 1607-11; see also Resnik, supra note 13, at 178-80; Hoffman & Berg, supra note 13, at 403-06.
118. Resnik, supra note 13, at 179.
119. Id.
120. Id.
121. Hoffman & Berg, supra note 13, at 411.
IRB members who are “highly dedicated to protecting human subjects and advancing biomedical research will not be deterred”\textsuperscript{123} by exposure to potential liability.

2. Viable Alternatives to the Legal System

Those against IRB liability emphasize that the legal system’s deterrent effect is overstated.\textsuperscript{124} They also point out that there are viable alternatives to the tort system. For example, some assert that moral values provide a much stronger incentive for appropriate behavior than does tort law, especially when IRB members “typically volunteer their services and know that they are charged with the duty of safeguarding the welfare of human subjects.”\textsuperscript{125} However, simply relying on IRB members’ good consciences is not working. Given the recent reprimands of certain IRBs, this reliance on individuals’ moral principles does not appear to be sufficient to promote attentive analyses by IRBs.\textsuperscript{126}

Critics against IRB liability also assert that an IRB accreditation system could be imposed as an alternative to litigation.\textsuperscript{127} An accrediting system would involve experts and peers developing a set of professional standards and criteria to govern IRB evaluations.\textsuperscript{128} The emphasis of accreditation would be on education and assuring that appropriate procedures are in place.\textsuperscript{129} Nonetheless, an accreditation system has inherent risks. First, accreditation programs are generally voluntary and represent a profession’s desire to self-regulate.\textsuperscript{130} The historical antecedents of the Tuskegee Syphilis Study, the Willowbrook Study, and the Jewish Chronic Disease Hospital Study demonstrate that abuses occur in self-regulated contexts. These same self-regulation concerns apply to IRBs that operate without sufficient governmental oversight. Moreover, there are concerns that IRB accreditation might add another layer of bureaucracy to an already heavily regulated environment.\textsuperscript{131} However, the legal system requires no further regulation or added bureaucracy. The tort system is a successful scheme already in place, and it can serve to curtail negligent IRB activities thereby resulting in more protection for human subjects who participate in research.

\textsuperscript{122} Steven E. Pegalis, A Proposal to Use Common Ground that Exists Between the Medical and Legal Professions to Promote a Culture of Safety, 51 N.Y.L.SCH. L. REV. 1056, 1060 (2006-2007).
\textsuperscript{123} Resnik, supra note 13, at 179.
\textsuperscript{124} E.g., Hoffman & Berg, supra note 13, at 408.
\textsuperscript{125} Id. at 408-09.
\textsuperscript{126} See discussion supra Section II.C.
\textsuperscript{127} Stalcup, supra note 3, at 1617; see also Resnik, supra note 13, at 135.
\textsuperscript{129} Id.
\textsuperscript{130} Id.
\textsuperscript{131} Id.
3. Defensive Practices

Critics also raise another concern. They argue that if IRBs are exposed to potential liability, then that will result in IRBs engaging in defensive practices in order to minimize their legal risks.  Some defensive IRB actions might include needless disapproval of research, requiring researchers to unnecessarily modify the research protocol or informed consent document, and increasing the frequency of its oversight functions. Few “would deny that an IRB should vigilantly protect human subjects.” In doing so, the IRB must carefully balance competing interests in research and make responsible decisions considering the totality of the circumstances. The key is for IRBs to find the appropriate balance between too little and too much protection. The precedents from IRB litigation could serve to clarify the appropriate standards to which IRBs must comply.

4. Increased Research Costs

Finally, critics assert that potential IRB litigation will raise the costs of research and development. They argue that costs will be generated by pre-trial and trial actions, as well as “by the need to provide IRBs with comprehensive [liability] insurance, to the extent that this is not already done.” However, the costs for IRB liability will likely only be a small percentage of the total cost of the research itself. Accordingly, research institutions will be able to spread the costs associated with liability across all studies done at the institution by factoring them into overhead charges.

B. Arguments in Favor of IRB Liability

Several arguments suggest that increased legal challenges to IRB activity will yield desirable outcomes. First, the risk of liability will promote more careful reviews by IRBs and more responsible decision-making. Second, the tort system serves as an external constraint on otherwise unfettered IRB power. Finally, the attention generated from IRB liability may lead to positive changes in the regulatory oversight of IRBs.

132. Resnik, supra note 13, at 179.
133. See id.
134. Id. at 180.
135. See id.
136. Id.
137. Hoffman & Berg, supra note 13, at 410.
138. Id. at 405.
139. Id. at 405-06.
1. Better IRB Decision-Making

Unfortunately, governmental oversight of IRBs is “extremely deficient to almost non-existent.” Governmental agencies lack the requisite amount of resources required to evaluate the work of thousands of IRBs. For instance, between June 1998 and March 2000, the Office for Protection from Research Risks (now the Office for Human Research Protections (OHRP)) conducted investigations at ten institutions and off-site document reviews at an additional 140 research entities. The Office found performance problems at a number of these institutions and required seven to suspend part of all of their federally funded research. As a result, IRBs have been left to regulate themselves. Recent incidents, however, demonstrate that IRB self-regulation is not sufficient to adequately protect human subjects.

In 1998, the OIG of the Department of Health and Human Services issued a report from its systematic, one-year study of IRBs which essentially identified an “IRB crisis.” Only a year later, the IRB at the University of Pennsylvania’s Gene Therapy Institute was cited by the OHRP for deficiencies in its approval of the gene therapy trial that resulted in Jesse Gelsinger’s death. Shortly thereafter in 2001, the OHRP found numerous inadequacies in the asthma study that caused the death of a healthy, twenty-four year old named Ellen Roche. In that same year, the Maryland Court of Appeals chimed in with its own IRB reprimand because it “abdicated its responsibility” in the Grimes lead paint abatement study. These examples aptly demonstrate why IRB self-regulation is not working. However, exposure to liability will lead to increased accountability for the IRB’s actions because the law contemplates that civil suits act as a warning that the law demands the exercise of due care. Therefore, potential liability will encourage independent, thorough, and objective decision-making.

2. Legal System as an Additional Constraint on IRB Power

The legal system is an external factor that is presumed to have a deterrent effect on human behavior. The tort system in particular operates under the

140. Id. at 407.
141. Id.
143. Id.
144. See discussion supra Section II.C.
145. See OFFICE OF INSPECTOR GEN., supra note 2.
146. See Letter from K.C. Borror to N. Nathanson, supra note 51.
149. See Pegalis, supra note 122, at 1060.
assumption that exposure to liability promotes public welfare because it makes “injury producing activities or goods more expensive than their safer competitors.”

The deterrent effect of liability is particularly important in the IRB context because the potential tort victims, injured human research subjects, have no control over the selection of IRB members. The deterrent effect of liability is further needed in the IRB context because, often times, unpaid IRB members make decisions based solely on their moral principles since there are no market forces operating to incentivize diligent and thorough reviews. Given that there is no economic factor present to induce quality IRB work, the legal system is a much needed check on otherwise unrestrained IRB power.

3. Positive Changes in the Regulatory Oversight of IRBs

Lastly, IRB liability should be encouraged because it serves as a vehicle for “broader formal changes in the law, regulatory activity, and institutional structures and behavior.” Litigation promotes public dialogue concerning important social issues; high profile cases will “capture public attention, raise awareness of oversight problems, and generate open debate concerning the regulatory standards.” For example, the occurrence of the Tuskegee Syphilis Study became known to the public in 1972. Shortly thereafter, the study was ordered to be shut down, and the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research issued the Belmont Report, which became the foundation for the human subject research regulations found in Title 45, Part 46 of the Code of Federal Regulations. Similar public exposure of IRB litigation could, in turn, lead to positive changes that protect research participants. Additionally, the public attention gained from IRB litigation might be accompanied by a willingness of institutions to invest more time and resources into the administration and training of IRB members.

For all these reasons, exposing IRBs to liability supports good public policy.

V. CONCLUSION

The issue of IRB liability has been justifiably magnified. The overwhelming number of protocols that IRBs review each year, with little federal agency...
oversight, has resulted in a system where IRBs are left to regulate themselves. Merely relying on the ethical judgments of IRB members, however, has not curtailed the frequency or severity of United States biomedical research tragedies. As a result, we essentially have an “IRB crisis” in the United States.

Our civil tort system adequately and appropriately provides a solution to our “IRB crisis” because the “system works in the context of rights and responsibilities.” In this respect, lawsuits naming an IRB or its members as defendants are desirable outcomes because they will promote more careful reviews and more responsible decision-making by IRBs, will serve as a check against IRBs’ unrestrained power, and will generate public attention likely leading to positive changes in the regulatory oversight of IRBs.

In light of the examples of Jesse Gelsinger, Ellen Roche, and the Grimes case, we can no longer allow IRB members to escape the consequences of misconduct. There is just too much at stake.

159. See id. at 407.
160. See discussion supra Section II.C.
161. Pegalis, supra note 122, at 1059.
THE DOCTRINE OF OTHER WRONGS:
A FRAMEWORK FOR PUNISHING CIVIL RECIDIVISM

Miriam R. Coles

Table of Contents
INTRODUCTION ............................................................................................68
I. CASE LAW ANALYSIS ........................................................................71
   A. Overview ....................................................................................71
   B. “Repeated Conduct” Language .................................................74
   C. “Potential Harm” Language .......................................................77
   D. Philip Morris – Example of Confusion ..................................78
II. REPREHENSIBILITY AND THE FIVE CATEGORIES OF OTHER
    WRONGS ............................................................................................79
    A. Harm to Others ........................................................................80
       1. Actual Harm to Others ......................................................81
       2. Potential Harm to Others ..................................................83
    B. Repeated Conduct ....................................................................87
       1. Repeated Acts ...................................................................88
       2. Nonadjudicated Repeated Conduct ..................................89
       3. Adjudicated Repeated Conduct ........................................91
III. RECIDIVISM ......................................................................................92
    A. Recidivism Background ............................................................92
    B. The Court’s Use of Recidivism and why it is Insufficient ............93
    C. Civil Recidivism Enhancements - Extraterritoriality and
       Dissimilar Conduct .....................................................................96
IV: PROBLEMS WITH MY PRESCRIPTION .................................................98
CONCLUSION .............................................................................................100
APPENDIX: INSTRUCTING THE JURY ON REPREHENSIBILITY ...............100

* Ms. Coles is a civil litigation attorney at Henry, Buchanan, Hudson, Suber & Carter P.A., located in Tallahassee, Florida, specializing in employment discrimination, civil rights defense, and general commercial tort litigation. Ms. Coles graduated Magna Cum Laude from Florida State University, College of Law in May of 2008.
INTRODUCTION

In 2000, the Oregonian widow of a heavy cigarette smoker sued the tobacco giant Philip Morris USA for negligence and deceit after her husband died of a smoking-related medical condition. The jury held Philip Morris liable on both counts, finding that the company intentionally deceived Williams and others by engaging in a forty year long, nationwide publicity campaign designed to create controversy over the health effects of smoking, despite the fact that no scientific controversy existed. The jury awarded $821,485 in compensatory damages and $79.5 million in punitive damages.

In upholding the amount of punitive damages, the Oregon Court of Appeals noted that the jury could have inferred from the evidence that the defendant’s conduct “caused a significant number of deaths each year in Oregon during the pertinent time period, together with other serious but nonfatal health problems . . . .” The court thus held that it was appropriate for the jury to “consider the effects of defendant’s actions on persons other than Williams in determining the amount of punitive damages.” The Oregon State Supreme Court agreed.

Throughout the five year appeal process, the jury instructions remained a central issue. At trial, Philip Morris proposed a set of jury instructions that specifically stated that although the jury could consider harm to others in determining the reprehensibility of the defendant’s conduct, the jury could not punish the defendant for the impact of harm to others. The Oregon Supreme Court rejected the proposed instruction and asserted that the “jury could consider whether Williams and his misfortune were merely exemplars of the harm that Philip Morris was prepared to inflict on the smoking public at large.” In addition, the court noted: “It is unclear to us how a jury could ‘consider’ harm to others, yet withhold that consideration from the punishment calculus. If a jury cannot punish for the conduct, then it is difficult to see why it may consider it at all.”

3. Id.
4. Id. at 839.
5. Id. at 841.
7. Id. at 1175 (“That instruction stated, in part: ‘The size of any punishment should bear a reasonable relationship to the harm caused to Jesse Williams by the defendant’s punishable misconduct. Although you may consider the extent of harm suffered by others in determining what that reasonable relationship is, you are not to punish the defendant for the impact of its alleged misconduct on other persons, who may bring lawsuits of their own in which other juries then resolve their claims and award punitive damages for those harms, as such other juries see fit.’”).
8. Id.
9. Id. at 1175 n.3.
Nevertheless, on appeal, the United States Supreme Court held that this is precisely what due process requires the jury to do. For the first time, the Court specifically ruled that a jury cannot punish a defendant for the harm its conduct inflicts upon “strangers to the litigation.” However, the Court clarified that while the jury cannot punish for harm to others, “evidence of actual harm to nonparties” is appropriate for determining the reprehensibility of a defendant’s conduct.

Justice Stevens observed the problem with this legal standard in his dissent. He stated: “When a jury increases a punitive damages award because injuries to third parties enhanced the reprehensibility of the defendant’s conduct, the jury is by definition punishing the defendant—directly—for third-party harm.” Legal scholars and judges have echoed Justice Stevens’ confusion over the Court’s holding, noting that such a standard will be practically impossible for a jury to understand, much less apply. In fact, lower courts began muddling the doctrine almost immediately.

A thorough reading of the Supreme Court cases leading up to Philip Morris sheds light on the confusion. First, the Court’s insistence that reprehensibility can be considered separate from punishment is, in this author’s opinion, impossible to comprehend. Reprehensibility, as the Court itself has noted, is the most important factor when assessing punitive damages. When the reprehensibility of the defendant increases, so too does the award for punitive damages. Therefore, reprehensibility is directly related to, and cannot be separated from, punishment.

11. Id.
12. Id. at 1064.
13. See id. at 1065-67 (Stevens, J., dissenting).
14. Id. at 1067 (Stevens, J., dissenting).
16. See, e.g., Moody v. Ford Motor Co., 506 F. Supp. 2d 823, 848-49 (N.D. Okla. 2007). The court incorrectly stated that “[a]lthough the jury may consider potential harm when determining the reprehensibility of the defendant’s conduct, this means only that the jury may consider the potential harm to the plaintiff, not all other potential victims . . . .” Id. The court later contradicted itself, correctly stating that “[u]nder Philip Morris, a jury may consider the fact that the defendant’s conduct is likely to harm third parties as an indication of reprehensibility, but the jury must be instructed not to award damages for actual or potential injuries to third parties.” Id. (emphasis added)). Here, the court confused punishment in the first sentence with the reprehensibility analysis. It is unclear whether this is an unintended oversight or if the punishment versus reprehensibility “nuance eludes” the court, as Justice Stevens would put it. See generally Philip Morris USA, 549 U.S. 346, 360 (2007) (Stevens, J., dissenting).
17. See Philip Morris USA, 549 U.S. at 355.
Another flaw in the *Philip Morris* decision is the Court’s failure to articulate a workable punitive damages “other wrongs” doctrine, or to correctly identify the types of harms to nonparties that should be considered when awarding punitive damages. Instead, the Court uses multiple terms interchangeably, supported by the same rationales, as though these multiple concepts describe the same type of harm or conduct. The result is a body of law that is grossly confusing and oversimplified.

Despite the Court’s failure to distinguish them, all other wrongs are not the same. In fact, there are seven factors lurking in the Court’s case law that distinguish other wrongs: (1) whether the wrong occurred during the same course of conduct that injured the plaintiff or a separate course of conduct, (2) whether the defendant committed the wrong against the plaintiff or a nonparty, (3) whether the wrong embodies actual or potential harm, (4) whether the wrong is adjudicated or nonadjudicated, (5) whether the wrong is intrajurisdictional or extraterritorial, (6) whether the wrong consists of a single or repeated acts, and (7) whether the wrong is similar to, or dissimilar from, the wrong of which the defendant is accused. Various combinations of these factors result in distinctive categories of other wrongs. This paper does not attempt to explore every possible combination of factors. Instead, it focuses on the combinations of other wrongs which the Supreme Court has inadvertently identified in its punitive damages opinions.

My contention is that, beneath the surface, the Supreme Court’s punitive damage case law contains five distinct categories of other wrongs. These categories are: (1) Actual Harm to Others caused by the same course of conduct that injured the plaintiff (2) Potential Harm/Risk of Harm to others (3) Repeated Acts against the plaintiff within the same course of conduct (4)  

18. The term “other wrongs” does not appear in any of the case law or punitive damages literature, nor has the Court specifically acknowledged this concept as a distinct doctrine of punitive damage law. Rather, the term was inspired by Federal Rule of Evidence 404(b), entitled “Other crimes, wrongs, or acts,” and seemed appropriate to describe what this author has pieced together from the Court’s punitive damage decisions as an implied doctrine. For the purposes of this paper, “other wrongs” is a general term including any evidence of other acts, crimes, conduct or harm committed by the defendant against the plaintiff and/or nonparties. This author struggled with an appropriate categorization for such evidence, convinced that terms like “harm to others” and “repeated conduct” were inadequate to describe all of the categories of evidence discussed in this paper. “Other wrongs,” on the other hand, seems broad enough to include all types of harm and conduct.


20. See *TXO Prod. Corp. v. Alliance Res. Corp.*, 509 U.S. 443, 460 (1993) (stating that it is “appropriate to consider the magnitude of the potential harm that the defendant's conduct would have caused to its intended victim if the wrongful plan had succeeded, as well as the possible harm to other victims that might have resulted if similar future behavior were not deterred.”); see also *Philip Morris USA*, 549 U.S. at 355-56 (discussing the relationship between “risk of harm” and the reprehensibility analysis). See also infra Parts I(C), I(D), II(A)(2).

Nonadjudicated Repeated Conduct,22 and (5) Adjudicated Repeated Conduct.23 Of these five categories, only Repeated Acts and Adjudicated Repeated Conduct should be considered when determining reprehensibility.24 Although the Court fails to recognize these as separate categories, they are not interchangeable. Rather, each category represents a unique type of harm and warrants individual consideration.

By isolating each category of other wrongs identified in the Court’s opinions and analyzing the appropriateness of each category for punishing the defendant and for determining reprehensibility, one can arrive at a punishment scheme that satisfies due process, accomplishes the goals of punitive damages, and is equitable to both parties. Just as importantly, one can devise a jury instruction that conveys the substantive limitations of due process in a manner that jurors can understand.25

Part I of this article analyzes the Supreme Court’s punitive damage case law addressing other wrongs. Part II explores the five categories of other wrongs and attempts to explain why some are appropriate for the reprehensibility analysis while others are not. Part III explains the need for a legitimate civil recidivism framework and describes how such a framework should operate; and, Part IV discusses the problems my prescription presents.

I. CASE LAW ANALYSIS

A. Overview

Philip Morris USA v. Williams, decided in early 2007, is the latest in a series of opinions addressing evidence of other wrongs. Each opinion mentions the concept in greater detail, and arguably more confusion, than the case before. Analyzing the evolution of the Court’s other wrongs doctrine sheds light on what the Court’s doctrine means, and also, on what it misses.

22. The term “nonadjudicated repeated conduct” does not appear in the case law. See infra Part II(B)(2) for a discussion of the origin and appropriateness of this term. See generally infra Part I(B) for an analysis of “repeated conduct” language in the case law.

23. The term “adjudicated repeated conduct” does not appear in the case law. See generally infra Part II(B)(3) and Part IV for a thorough discussion of the origin and appropriateness of this term.


25. See Appendix outlining my model jury instruction, which fully incorporates my prescription.
In *Pacific Mutual Life Insurance Co. v. Haslip*, decided in 1991, the Court held for the first time that the Due Process Clause places procedural due process requirements on punitive damage awards.\(^{26}\) In *Haslip*, plaintiffs sued their insurance company, Pacific Mutual Life Insurance, because its agent embezzled their insurance premiums, causing their policy to lapse.\(^{27}\) The jury awarded the plaintiff’s $200,000 in compensatory damages and $840,000 in punitive damages.\(^{28}\) The Supreme Court upheld the jury award, ruling that the procedures Alabama used to review punitive damages awards satisfied due process.\(^{29}\)

Just two years later, in *TXO Production Corp. v. Alliance Resources Corp.*, the Court recognized a substantive due process limitation on the actual size of awards and created a “reasonableness test” for assessing the constitutionality of punitive damage awards.\(^{30}\) In *TXO*, an oil and gas development project, TXO Production Co., brought action against Alliance Resources to clear an allegedly clouded title.\(^{31}\) Alliance Resources counterclaimed for slander of title and alleged that TXO had filed a bad faith lawsuit solely for the purpose of strong-arming Alliance into renegotiating a royalty deal.\(^{32}\) After hearing evidence, including testimony that TXO stood to gain as much as $37 million from its scheme, the jury awarded Alliance $19,000 in compensatory damages and $10 million in punitive damages.\(^{33}\)

On appeal, TXO argued that because the punitive damage award was 526 times greater than the amount of compensatory damages, the award was unconstitutionally arbitrary and excessive.\(^{34}\) The Court held that although high ratios between the compensatory and punitive damage awards are suspect, they can be justified by other factors, such as the potential harm to the defendant and to others.\(^{35}\) The court stated that when determining the:

relationship between actual and punitive damages . . . it is appropriate to consider the magnitude of the potential harm that the defendant's conduct would have caused to its intended victim if the wrongful plan had succeeded, as well as the possible harm to other victims that might have resulted if similar future behavior was not deterred.\(^{36}\)


\(^{28}\) Id. at 7 n.2.

\(^{29}\) Id. at 23-24.


\(^{32}\) Id. at 450.

\(^{33}\) Id. at 451.

\(^{34}\) Id. at 453.

\(^{35}\) Id. at 460.

\(^{36}\) Id.
In *BMW of North America v. Gore*, the Court clarified the “reasonableness test” presented in *TXO* and issued concrete guideposts for making substantive assessments of punitive damage awards.\(^{37}\) In *BMW*, a customer sued BMW for fraud after the company failed to disclose pre-sale damage repair, which had depreciated the value of the customer’s car by approximately ten percent.\(^{38}\) The jury awarded $4,000 in compensatory damages and $4 million in punitive damages.\(^{39}\) The Court held that the punitive damage award violated substantive due process because the jury based the award on evidence of lawful, out-of-state conduct.\(^{40}\)

In addition, the Court clarified the “reasonableness test” presented in *TXO*.\(^{41}\) First, the Court stated that the award cannot be grossly excessive in comparison to the state’s legitimate interest in punishment and deterrence.\(^{42}\) Then, the award should be assessed based on three guideposts: (1) the degree of reprehensibility (2) the disparity between the harm or potential harm suffered by the plaintiff and the punitive damage award, and (3) the difference between the punitive damage award and the civil and criminal penalties authorized for such conduct.\(^{43}\)

The Court also elaborated on guidepost one, stating that reprehensibility is the most important factor in assessing a punitive damage amount,\(^{44}\) and that reprehensibility should be viewed as a:

| Sliding scale of reprehensible conduct: violent crimes more reprehensible than nonviolent ones, trickery and deceit more reprehensible than negligence, conduct causing physical harm more reprehensible than conduct causing purely economic harm, deliberate false statements more reprehensible than omissions of material facts, and repeated conduct more reprehensible than an isolated incident.\(^{45}\) |

In 2003, the Court ruled again on the issues of substantive and procedural due process in *State Farm Mutual Automobile Insurance Co. v. Campbell*.\(^{46}\) In *State Farm*, after the plaintiffs caused a serious automobile accident, they sued their car insurance company, State Farm, for a bad-faith failure to settle a claim within policy limits, fraud, as well as intentional infliction of emotional distress.\(^{37}\) The jury awarded $2.6 million in compensatory damages and $145 million in punitive damages.\(^{48}\) The Court held the jury’s $145 million award

---

38. Id. at 563.
39. Id. at 564.
40. See id. at 573-74.
41. See id. at 568.
42. Id.
44. Id.
45. Franze & Scheuerman, supra note 26, at 456 (citing *Gore*, 517 U.S. at 575-76).
47. Id. at 414.
48. Id. at 415.
unconstitutionally excessive.\textsuperscript{49} The Court ruled that the jury could not punish the defendant for lawful or unlawful conduct occurring in another state\textsuperscript{50} and suggested that procedurally, juries must be instructed on certain substantive due process limitations.\textsuperscript{51} However, the Court limited this holding, stating that lawful out-of-state conduct can be used for a proper purpose such as demonstrating “deliberateness or culpability” if the conduct has a sufficient “nexus to the specific harm suffered by the plaintiff.”\textsuperscript{52}

In \textit{State Farm v. Campbell}, the Court specifically enumerated the reprehensibility factors mentioned in \textit{BMW v. Gore}.\textsuperscript{53} The Court stated:

\begin{quote}
We have instructed courts to determine the reprehensibility of a defendant by considering whether: the harm caused was physical as opposed to economic; the tortious conduct evinced an indifference to or a reckless disregard of the health or safety of others; the target of the conduct had financial vulnerability; the conduct involved repeated actions (emphasis added) or was an isolated incident; and the harm was the result of intentional malice, trickery, or deceit, or mere accident.\textsuperscript{54}
\end{quote}

The inclusion of “repeated actions” in the \textit{State Farm} reprehensibility factors likely served as precedent for the Court’s continued holding in \textit{Philip Morris} that evidence of actual harm to others can be considered when determining reprehensibility.\textsuperscript{55} However, a closer look at the progression of other wrongs language in \textit{Pacific Mutual v. Haslip}, \textit{TXO v. Alliance}, \textit{BMW v. Gore}, \textit{State Farm v. Campbell}, and \textit{Philip Morris USA v. Williams} casts doubt on the appropriateness of this association.

\textbf{B. “Repeated Conduct” Language}

In the first case, \textit{Pacific Mutual v. Haslip}, Justice Blackmun’s majority opinion made no mention of actual or potential harm to others. However, it

\textsuperscript{49} Id. at 429.
\textsuperscript{50} Id. at 421.
\textsuperscript{51} Id. at 422.
\textsuperscript{52} State Farm Mut. Auto. Ins. Co. v. Campbell, 538 U.S. 408, 422 (2003) (holding that “lawful out-of-state conduct may be probative when it demonstrates the deliberateness and culpability of the defendant’s action in the State where it is tortious, but that conduct must have a nexus to the specific harm suffered by the plaintiff. A jury must be instructed, furthermore, that it may not use evidence of out-of-state conduct to punish a defendant for action that was lawful in the jurisdiction where it occurred”). It is unclear whether the Court means that evidence of lawful out-of-state conduct can be introduced solely for determining reprehensibility or if it can only be introduced when necessary to show intent, motive, plan, etc. during the plaintiff’s case-in-chief.
\textsuperscript{53} Id. at 419 (citing BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 575 (1996)).
\textsuperscript{54} Id.
\textsuperscript{55} See Philip Morris USA v. Williams, 549 U.S. 346, 355 (2007). This proposition is based largely on my analysis of the case law, as discussed in this Part, and not on the text of the opinion. Justice Breyer, writing the majority opinion in \textit{Philip Morris}, did not cite any authority for the assertion that the jury should hear evidence of actual harm to others when determining reprehensibility. See id.
indirectly mentioned the concept of repeated conduct when approving of the Alabama Appeals Court’s usage of “the existence and frequency of similar past conduct” (emphasis added) as a factor for determining reprehensibility.\(^{56}\) Although the opinion did not explain the meaning or scope of this factor, it is continually cited in the Court’s subsequent opinions.\(^{57}\) In fact, in later cases, the Court cited to this passage when discussing recidivism in the civil context.\(^{58}\)

In *TXO v. Alliance*, the Court cited “the existence and frequency of similar past conduct” from *Haslip* in support of the proposition that a “pattern and practice of impermissible behavior” is a relevant factor for increasing the size of a punitive damage award.\(^{59}\)

In *Gore*, the Court determined that repeated conduct is more reprehensible than an isolated incident.\(^{60}\) To support this assertion, Justice Stevens, writing the majority opinion, cited to the “pattern and practice” passage in *TXO*, which in turn cited to *Haslip’s “existence and frequency of similar past conduct.”*\(^{61}\)

Although *Gore* is cited in *State Farm* as expressly instructing courts to consider five specific factors when determining the level of reprehensibility,\(^{62}\) the *Gore* Court’s “instruction” is much less formal. First, it must be understood that *Gore* was the Court’s first attempt to clearly distinguish punishment and punitive damage ratios from the reprehensibility analysis.\(^{63}\) In addition, the discussion of reprehensibility in *Gore* did not enumerate specific factors. Rather, the *State Farm* Court gleaned five factors, including “repeated acts” from the two-page discussion of reprehensibility in *Gore*.\(^{64}\) Notably, in this discussion of

---

58. *State Farm*, 538 U.S. at 423 (“Although [o]ur holdings that a recidivist may be punished more severely than a first offender recognize that repeated conduct is more reprehensible than an individual instance of malfeasance, *Gore [517 U.S.] at 577, in the context of civil actions courts must ensure that the conduct in question replicates the prior transgressions. *TXO*, 509 U.S. at 462 n.28.”); *Gore*, 517 U.S. at 576-77 (“Evidence that a defendant has repeatedly engaged in prohibited conduct while knowing or suspecting that it was unlawful would provide relevant support for an argument that strong medicine is required to cure the defendant's disrespect for the law. *See* [TXO, 509 U.S.] at 462 n. 28. Our holdings that a recidivist may be punished more severely than a first offender recognize that repeated misconduct is more reprehensible than an individual instance of malfeasance.”).
60. *Gore*, 517 U.S. at 577.
61. *Id.* at 576.
reprehensibility, the Court cites a significant number of criminal cases dealing with the constitutionality of criminal punishments.65

In State Farm, the Court only listed two reprehensibility factors that potentially implicate harm to others: (1) the tortious conduct evinced an indifference or reckless disregard to the health and safety of others and (2) the conduct involved repeated acts or was an isolated incident.66 In support of these factors, the State Farm Court cited the two-page analysis of reprehensibility in Gore mentioned above.67 The section of that two-page analysis discussing harm to others states:

[E]vidence that a defendant has repeatedly engaged in prohibited conduct while knowing or suspecting that it was unlawful would provide relevant support for an argument that strong medicine is required to cure the defendant’s disrespect for the law. (citation omitted). Our holdings that a recidivist may be punished more severely than a first offender recognize that repeated misconduct is more reprehensible than an individual instance of malfeasance (citation omitted).68

As mentioned above, the Gore Court cited the “pattern and practice” passage in TXO, which in turn cited Haslip’s “existence and frequency of past similar conduct” factor.69 Hence, State Farm’s “conduct involv[ing] repeated actions,”70 came from Gore’s “repeated misconduct,”71 which came from TXO’s “pattern and practice,”72 which came from Haslip’s “existence and frequency of similar past conduct.”73

Logically, one can assume that if all of these passages relating to other wrongs cited the same underlying passage in Haslip, then the Court meant for them to represent the same idea. Consequently, because the Gore passage - supported by Haslip and TXO and used as the foundation for the reprehensibility factors enumerated in State Farm - is framed around recidivism,74 it is fairly certain that references to “repeated acts,”75 “repeated misconduct,”76 “similar

---

66. State Farm, 538 U.S. at 419.
67. Id.
68. Gore, 517 U.S. at 576-77.
72. TXO, 509 U.S. at 469.
74. See supra note 58 and accompanying text.
75. State Farm, 538 U.S. at 419.
past conduct,“77 and “a pattern and practice” of improper behavior78 are all justified, in the Court’s mind, by recidivism. Therefore, it is ultimately recidivism that impacts reprehensibility.

It is also significant that prior to Philip Morris the Court never mentioned “actual harm to others.” It seems relatively certain from the text of the Philip Morris opinion that although the Court used the two other wrongs terms, it did not intend to differentiate actual harm to others from repeated conduct.79 In fact, just as the Gore Court referenced recidivism as a justification for punishing repeated conduct more severely than isolated conduct80 the Philip Morris Court referenced recidivism to support its contention that “actual harm to others” is relevant to reprehensibility.81 Therefore, the Court likely takes the mistaken view that actual harm and repeated conduct represent the same concept. The Court believes that both categories are justified by principles of recidivism.

C. “Potential Harm” Language

Subsection B analyzed the Court’s use of multiple terms – repeated acts, repeated conduct, patterns and practice, and actual harm to others – when attempting to explain the doctrine of other wrongs. Unfortunately, the Court has further complicated this analysis with its references to “potential harm”82 and risk of harm.83

In TXO, the Court focused, not on punishment for repeated conduct, but on punishment for the potential harm the defendant’s conduct might have caused.84 The Court held that evidence of potential harm to the defendant and those similarly situated is appropriate for determining whether or not the relationship between the harm and the punitive damage award is reasonable.85

Although the State Farm Court limited the holding in TXO to allow only evidence of potential harm to the plaintiff when determining the harm to punishment ratio,86 the Philip Morris Court affirmed potential harm to others as a reprehensibility factor.87 In addition, the State Farm opinion also endorsed potential harm to others by including it in one of the five reprehensibility factors – whether “the tortious conduct evinced an indifference to or a reckless disregard

76. Gore, 517 U.S. at 577.
77. Haslip, 499 U.S. at 21.
79. See infra Part I(D).
82. TXO, 509 U.S. at 460.
83. See Philip Morris USA, 549 U.S. at 357.
85. Id.
of the health or safety of others.” Interestingly, this potential harm passage in *State Farm* is supported by a citation to *Gore*’s discussion of “repeated misconduct,” which in turn cited *TXO*’s discussion of “pattern and practice,” which cited *Haslip*’s “existence of frequency or similar past conduct” factor.

### D. Philip Morris – Example of Confusion

One of the main problems with the case law mentioned in Part I is that the Court fails to differentiate between the five categories of other wrongs. The following passage from *Philip Morris* illustrates the confusion caused by the Court’s interchangeable use of other wrongs language:

> We did not previously hold explicitly that a jury may not punish for the harm caused others (emphasis added). But we do so hold now . . . We have explained why we believe the Due Process Clause prohibits a State’s inflicting punishment for harm caused (emphasis added) strangers to the litigation. At the same time we recognize that conduct that risks harm (emphasis added) to many is likely more reprehensible than conduct that risks harm (emphasis added) to only a few. And a jury consequently may take this fact into account in determining reprehensibility. Cf., e.g., *Witte v. United States*, 515 U.S. 389, 400 (1995) (recidivism statutes taking into account a criminal defendant’s other misconduct (emphasis added) do not impose an “‘additional penalty for the earlier crimes,’ but instead . . . ‘a stiffened penalty for the latest crime, which is considered to be an aggravated offense because a repetitive one.’”).

In this passage, the Court begins by discussing actual harm caused to nonparties. Then, without clarification, the Court explains that while the jury cannot punish for harm caused in the past, the risk of future harm is relevant to the determination of reprehensibility. Does this mean that the court intends evidence of actual and potential harm to be analyzed in the same manner? The Court does not explain whether or not it is making a distinction between actual harm and potential harm. Given the fact that most of the opinion discusses actual harm to others, the Court is probably using the concepts interchangeably, saying that any actual or potential harm to nonparties cannot be considered for punishment but can be considered for determining reprehensibility. However, the Court could be saying that while actual harm caused to nonparties is not relevant to reprehensibility, a risk of harm to many is relevant.

---

88. *State Farm*, 538 U.S. at 419.
89. *Id.* (citing BMW of N. Am., Inc. v. *Gore*, 517 U.S. 559, 576-77 (1996)).
93. *See id.*
94. *Id.*
The distinction between the types of harm mentioned in this passage is further clouded by the Court’s use of recidivism statutes to support the assertion that conduct that risks harm to many is more reprehensible than conduct that risks harm to only a few. If the Court actually means risk of harm as a distinct category of conduct in this sentence, then recidivism is inapplicable. Recidivism statutes increase punishment for criminal conduct that has already occurred, not for conduct that risks harm in the future.\(^{95}\) In addition, in both *Gore* and *State Farm*, the Court used recidivism to discuss past repeated conduct, not risk of harm.\(^{96}\)

In summary, the case law discussing the relationship between other wrongs and reprehensibility is confusing. The Court uses multiple terms to describe what it likely believes to be the same concept. However, amidst this confusion, two prevalent themes emerge: repeated conduct and recidivism. Ultimately, all types of other wrongs can be traced back to repeated conduct and, consequently, the primary justification for considering repeated conduct when determining reprehensibility is recidivism.

**II. REPREHENSIBILITY AND THE FIVE CATEGORIES OF OTHER WRONGS**

Although the Court uses other wrongs language interchangeably, it seems to have unintentionally identified three major groups: repeated conduct,\(^{97}\) potential harm to others,\(^{98}\) and actual harm to others.\(^{99}\) In order for the other wrongs analysis to make logical, ethical, and legal sense, the Court should acknowledge that these categories represent different concepts and treat each category distinctly for purposes of determining reprehensibility. In addition, the Court should separate repeated conduct into three distinguishable categories: “repeated acts,”\(^{100}\) “nonadjudicated repeated conduct,”\(^{101}\) and “adjudicated repeated conduct.”\(^{102}\) Of these five categories of evidence, only repeated acts and adjudicated repeated conduct should be considered during the reprehensibility analysis.

When assessing punitive damages, the five categories of other wrongs should be treated as follows:

(1) **Actual Harm to Others:** As the Court ruled in *Philip Morris*, considering actual harm to others is inappropriate for punishment.\(^{103}\) However, contrary to the Court’s holding, it is also inappropriate for consideration when determining

---

95. See infra Part III.
96. See supra note 58 and accompanying text.
97. See supra Part I(B).
98. See supra Part I(C).
99. See supra Part I(B).
100. See infra Part II(B)(1).
101. See infra Part II(B)(2).
102. See infra Part II(B)(3).
reprehensibility. Actual harm to others should only be considered when determining the scope of the defendant’s conduct and whether or not it was malicious or wanton and reckless. 104

(2) Potential Harm/Risk of Harm to Others: Evidence of potential harm to others is not appropriate for determining the amount of punishment. 105 But, the Court currently allows juries to consider evidence of potential harm to others when determining reprehensibility. 106 Arguably, potential harm to others does impact the reprehensibility of a defendant’s conduct. However, there are multiple reasons why evidence of risk of harm to others should be excluded from the reprehensibility analysis. 107

(3) Repeated Acts: The defendant’s repeated acts against the plaintiff during the same course of conduct are relevant to both reprehensibility and punishment. 108

(4) Nonadjudicated Repeated Conduct: The Court does not specifically discuss nonadjudicated repeated conduct. The rules of evidence dictate that nonadjudicated repeated conduct is not admissible for the purpose of punishment. Therefore, it should not be included in the reprehensibility analysis. However, this evidence is permissible for proper purposes, such as demonstrating intent, motive, plan, etc. 109

(5) Adjudicated Repeated Conduct: Evidence of adjudicated repeated conduct is appropriate for determining reprehensibility as part of a legitimate civil recidivism framework. 110

The remainder of Part II explores the five categories of other wrongs in detail, focusing on their appropriateness for consideration when determining reprehensibility.

A. Harm to Others

Punitive damages are linked to an individual plaintiff’s harm and his ability to recover damages for that harm in a civil suit. 111 Many legal scholars and judges insist that punitive damages are punishment for social wrongs. 112 As

104. See infra Part II(A)(1).
106. Philip Morris, 549 U.S. at 357.
107. See infra Part II(A)(2).
108. See infra Part II(B)(1).
109. See infra Part II(B)(2).
110. See infra Part II(B)(3) and Part III.
111. See Thomas B. Colby, Beyond the Multiple Punishment Problem: Punitive Damages as Punishment for Individual Private Wrongs, 87 MINN. L. REV. 583, 622-23 (2003).
112. See id. at 603-06; see also, e.g., Elizabeth J. Cabraser, Unfinished Business: Reaching the Due Process Limits of Punitive Damages in Tobacco Litigation Through Unitary Classwide Adjudication, 36 WAKE FOREST L. REV. 979, 981 (2001); Theodore B. Olson & Theodore J. Boutrous, Jr., Constitutional Restraints on the Doctrine of Punitive Damages, 17 PEPP. L. REV.
appealing as it may sound, this description of punitive damages fails to adequately describe either the history of punitive damages or current case law.\footnote{Colby, supra note 111, at 602-42; see also Anthony J. Sebok, Punitive Damages: From Myth to Theory, 92 IOWA L. REV. 957, 1006-29 (2007) (discussing punitive damages as personal punishment for the violation of a private right to respect).} In early punitive damages cases, courts upheld the validity of multiple punitive damage awards on the basis that each award only punished the defendant for the harm done to the plaintiff.\footnote{Colby, supra note 111, at 619. See, e.g., Ala. Power Co. v. Goodwin, 99 So. 158, 160 (Ala. 1923) (“In its civil aspects the single act or omission forms as many distinct and unrelated wrongs as there are individuals injured by it.”); Tullage v. Wade, 95 Eng. Rep. 909, 909 (C.B. 1769) (holding that it was permissible for a father and daughter to recover for the same act only because the jury on the father’s case was properly instructed not to punish for harm to the daughter and because the award was low enough in the father’s case to infer that the jury followed this instruction).} The courts rightly held that “a single illicit act that harms more than one person constitutes a number of legally distinct private wrongs.”\footnote{Colby, supra, note 111, at 622.}

Today, the Court’s rulings in \textit{BMW v. Gore}, \textit{Campbell v. State Farm}, and \textit{Philip Morris USA v. Williams} suggest that the Court still views punitive damages as a creature of individual tort law rather than a tool for social sanctions. Thus, while punitive damages may punish a defendant for conduct that juries consider worthy of society’s moral condemnation; that condemnation should be directly linked to the plaintiff’s harm.\footnote{See id. at 607.}

1. Actual Harm to Others

“Actual harm to others”\footnote{Id. at 607-08.} emerged for the first time in \textit{Philip Morris}, and in the Court’s mind, likely embodies the same concept as repeated conduct.\footnote{See supra Part I(B).} However, for my purposes, these factors are distinct. “Actual harm to others” should be limited to the concept of concurrent harm - harm caused to nonparties during the same course of conduct that injured the plaintiff. This concept should not include persons the defendant harmed in the past or present through a separate course of conduct. In addition, actual harm to others describes the harm that \textit{actually materializes} from a defendant’s specific course of conduct, irrespective of the defendant’s initial intent.

The jury should not consider evidence of actual harm to others when determining the reprehensibility of the underlying conduct. Conduct that actually harms many people is not in and of itself more reprehensible than conduct that only harms the plaintiff. The conduct itself is the same – the same plan, the same intent, the same action(s). Because the conduct is the same, the reprehensibility
of the conduct is also the same, regardless of how many people happen to be injured by the conduct.

In *Philip Morris*, for example, the defendant intentionally set out to defraud its consumers.\(^\text{119}\) If Philip Morris’ course of conduct had only harmed Jesse Williams, instead of a large number of Oregonians, the conduct would still have been the same. Philip Morris’ conduct embodied the same intent, plan, and actions regardless of how many people were actually harmed. Thus, the reprehensibility level should be determined by Philip Morris’ intended conduct, not by the unintended result of that conduct.

This does not mean that a defendant, like Philip Morris, should receive the same punishment no matter how many people its conduct harms. Instead, the amount of punishment the conduct deserves should be repeated for each person the defendant harmed.\(^\text{120}\) Using Philip Morris as an example, assume that the conduct deserves a reprehensibility of ten. If Philip Morris only harmed Williams, then the total punishment would be ten. However, if Philip Morris harmed 100 people, then the total punishment would be ten x 100. The fact that Philip Morris actually harmed multiple people increases the total punishment. But, although the overall punishment is greater, the underlying conduct itself is not more reprehensible simply because many people were injured.

In addition to the theoretical problems mentioned above, basing a reprehensibility determination on evidence of actual harm to others violates due process. In *Philip Morris*, the Court stated that due process forbids a jury to punish a defendant for harm to nonparties because “the Due Process Clause prohibits a State from punishing an individual without first providing that individual with an opportunity to present every available defense’ (citation omitted). Yet a defendant threatened with punishment for injuring a nonparty victim has no opportunity to defend against the charge . . . .”\(^\text{121}\) This same reasoning explains why the jury cannot increase reprehensibility based on harm to nonparties. Increasing reprehensibility increases the punishment. And, increasing punishment without giving the defendant a chance to defend against the alleged harm violates the principles of due process.\(^\text{122}\)

Actual harm to others should not be considered when determining reprehensibility because it does not increase culpability and it potentially

---

120. This will not be accomplished in the plaintiff’s suit. Rather, the total punishment will not be received until all of the victims bring their own lawsuits. This could be simplified through class actions.
122. See generally id. at 1066-67 (Stevens, J., dissenting) (arguing that considering third-party harm when assessing reprehensibility cannot be distinguished from considering it to assess punishment directly). If Stevens is correct, then allowing the jury to consider actual harm to others when determining reprehensibility violates due process for the same reason that considering actual harm to other when punishing the defendant violates due process.
violates due process. However, it may be introduced during phase one of the trial to show that the defendant’s conduct was malicious or wanton and reckless.123

2. Potential Harm to Others

Potential harm to others emerged in TXO and appears in the State Farm factor considering whether “the tortious conduct evinced an indifference to or a reckless disregard of the health or safety of others.”124 For the purposes of this paper, potential harm encompasses risk of harm to nonparties, which the defendant knew or should have known existed as it planned and implemented the course of conduct. Actual harm to others embodies outcomes, not intent. Potential harm, on the other hand, considers the defendant’s intended harm, rather than the harm that actually occurs. Unlike actual harm to others, which materializes after the conduct begins, potential harm to others, contemplated when the conduct is contrived and initiated, may be relevant to reprehensibility.

Some scholars argue that basing culpability on actual harm alone is irrational and does not serve the goals of punishment.125 If the defendant risks harm, he should not be shielded from liability because the harm did not materialize.126 Others argue that from a moral and retributive perspective, the person who brings about the harm is more culpable than the person who does not.127 From a purely moral standpoint, risk of harm, unlike actual harm, seems appropriate to the reprehensibility determination. Risk of harm speaks to the defendant’s mindset when he began the course of conduct at issue.

One example of the distinction between actual harm to others and risk of harm is the bomber hypo Justice Stevens used in his dissent to the majority’s opinion in Philip Morris.128 In support of his position that harm to nonparties should be considered when punishing a defendant, Stevens stated that “[a] murderer who kills his victim by throwing a bomb that injures dozens of

---

123. In most courts, trial judges have the option of bifurcating, or separating the issues for trial. “In a bifurcated trial, the jury hears evidence on one or more issues and decides those issues prior to hearing the evidence on any remaining issues.” In a case for punitive damages, the plaintiff has an absolute right to have the punitive damage award issue bifurcated from the issue of liability. Bifurcation is a procedural tool used to protect defendants from evidentiary prejudice. Steven S. Gensler, Prejudice, Confusion, and the Bifurcated Civil Jury Trial: Lessons from Tennessee, 67 TENN. L. REV. 653, 653-61, 674 (2000).
124. See supra Part I(C).
127. Id.
bystanders should be punished more severely than one who harms no one other than his intended victim.”

In this hypo, Stevens illustrates actual harm to nonparties. As presented, the hypo seems easy to refute. Both acts involved the same conduct and the same intent - to kill one person using a bomb. The conducts are equally reprehensible. The murderer who kills bystanders should receive a greater punishment, not because his conduct is more reprehensible, but because his sentence will reflect a punishment for each person killed. Here, actual harm is not relevant to the reprehensibility analysis.

However, altering the hypo slightly reveals the true moral dilemma. Imagine that murderer number two waits until no one is around before throwing his bomb because he is unwilling to risk harm to others besides his intended victim. Murderer number one, on the other hand, knows and does not care that throwing the bomb into the crowd risks harm to many others. It seems that murderer number one is more culpable because he created more potential harm than murderer number two. Thus, potential harm or risk of harm to others is arguably relevant to the reprehensibility analysis while actual harm to others is not.

The distinction between the relevance of actual and potential harm to reprehensibility is further exemplified by analyzing the following statement: “It may be reprehensible to falsely cry fire in a movie theatre with one person inside, but it is infinitely more reprehensible to do so when the theatre is full.” Shouting fire in a crowded theatre is only more reprehensible if the person knows how many people are in the theatre. Otherwise, both actions are equally reprehensible; however, the punishment level will differ based on the number of people actually harmed.

Applying this to Philip Morris, the fact that the company intentionally risked harm to a large number of people on an ongoing basis seems to make it more culpable. And, it is potentially relevant to establishing what exactly the course of conduct entailed.

On the other hand, considering risk of harm may be inappropriate for tort law. Although punitive damages are quasi-criminal in nature, they are a part of the tort system. Criminal law often provides for the punishment of potential harm, provided the defendant meets a certain mens rea requirement and takes a “substantial step toward completion” of the crime. Attempt charges,

130. Id.
preventative crimes like DUI and charge for conspiracy to commit crimes all punish for unrealized, potential harm.\textsuperscript{134}

But in tort law, a wrongdoer is only liable if the risk of harm he created is “realized” and causes actual harm.\textsuperscript{135} In most cases, our tort system does not provide a basis for plaintiffs to sue for misconduct absent a showing of harm.\textsuperscript{136}

Even the doctrine of nominal damages supports this assertion. Courts universally accept the rule that punitive damages are only appropriate when the plaintiff shows that he suffered actual damage.\textsuperscript{137} Some courts disallow punitive damages if the plaintiff is only entitled to nominal damages.\textsuperscript{138} These courts reason that “if the individual cannot show actual harm, he or she has but a nominal interest, hence, society has little interest in having the unlawful, but otherwise harmless, conduct deterred . . . .”\textsuperscript{139}

However, many jurisdictions permit punitive damage awards even when the plaintiff is only awarded nominal compensatory damages.\textsuperscript{140} These jurisdictions hold that as long as the plaintiff shows the invasion of a “legally protected interest,” a punitive damage award is appropriate.\textsuperscript{141}

Both views of nominal damages support the presumption that punitive damages are directly linked to the actual harm of the plaintiff. This actual harm does not have to be a traditional, compensatory harm. For example, even jurisdictions that normally reject awards based on nominal damages acknowledge that the violation of an intangible legal interest, such as the right to keep others off of your property, is an actual harm sufficient to allow a punitive award.\textsuperscript{142} Even when nominal damages are sufficient to trigger punitive damages, the court is punishing the defendant for some actual harm, be it tangible or intangible, that his conduct caused the plaintiff.

Some may argue that if the jury can consider the potential harm to the plaintiff when assessing punitive damages,\textsuperscript{143} it should be able to consider potential harm to others as well. However, the Court should require potential harm to be linked to the plaintiff’s actual harm in order for punitive damages to be justified.

\textsuperscript{134} See id. at 1636-37.
\textsuperscript{135} Id. at 1638.
\textsuperscript{137} Richard C. Tinney, Annotation, Sufficiency of Showing of Actual Damages to Support Award of Punitive Damages—Modern Cases, 40 A.L.R. 4th 11, 18 (1985).
\textsuperscript{138} 22 AM. JUR. 2D Damages § 554 (2007).
\textsuperscript{139} Jaque v. Steenberg Homes, Inc., 563 N.W.2d 154, 158 (Wis. 1997).
\textsuperscript{140} 22 AM. JUR. 2D Damages § 554 (2007).
\textsuperscript{141} Id. (citing Harwood v. Talbert, 39 P.3d 612 (Idaho 2001)).
\textsuperscript{142} Jaque, 563 N.W.2d at 158-59.
It is true that in the medical context, some courts have granted standing based on the plaintiff’s increased risk of harm, such as the risk of developing cancer after negligent exposure to toxins. In addition, some scholars argue that risk of harm is itself an actual harm because it reduces the person’s overall welfare, regardless of whether or not the harm occurs. However, in these contexts, the plaintiff is recovering for risk of harm to himself, not to nonparties. And, to recover, the plaintiff will have to prove that the defendant’s conduct caused the increased risk of harm beyond a preponderance of the evidence. Allowing a plaintiff to recover for risk of harm to himself is consistent with the idea of including potential harm to the plaintiff in the punitive damages punishment determination. However, it does not go so far as to justify allowing a plaintiff to recover punitive damages for risk of harm to nonparties. Just as potential harm to others is not appropriate for consideration when determining the amount of punishment, so too, the plaintiff should not be permitted to recover punitive damages based on a reprehensibility analysis that reflects risk of harm to nonparties.

Because punitive damages are linked to a showing of actual harm to the plaintiff, reprehensibility should only be linked to the plaintiff’s actual or potential harm. Otherwise, because the tort system does not punish for unrealized torts, awarding punitive damages based on reprehensibility that has been increased for potential harm to others will lead to unjust outcomes. Consider the following examples:

**Scenario One** - Philip Morris planned and began its fraudulent campaign, with which it is willing to harm 1,000 people. But, before anyone is actually harmed, the conduct is stopped. In this example, Philip Morris faces no civil liability under regular tort law because the conduct does not cause any actual harm.

**Scenario Two** - Philip Morris has planned and began its fraudulent campaign, with which it is willing to harm 1,000 people. But, after harming only 1 person (Williams) the conduct is stopped and no one else is harmed. Philip Morris is now liable to Williams for compensatory and punitive damages.

In scenario two, if the state considers the risk of harm to the 1,000 people when determining reprehensibility, it will be in effect, punishing the defendant for risking harm to others. Although the defendant’s conduct is the same in both scenarios, scenario one results in no punitive damages, while scenario two results in punitive damages determined by a reprehensibility analysis reflecting a potential risk of harm to others.

---

146. See *id.* at 979-83 (providing various examples where courts allowed a plaintiff to recover for increased risk of harm).
risk of harm to 1000 people. This is clearly inequitable. If actual harm to Williams is needed to trigger punishment, then the reprehensibility of the conduct should be viewed only in relation to Williams.

There is also a serious risk that the potential harm analysis can get out of hand. Consider, for example, *Action Marina Inc. v. Continental Carbon Inc.*, an Eleventh Circuit case decided after *Philip Morris*. In this case, the Court awarded 1.2 million dollars in compensatory damages against a manufacturer for intentionally damaging the plaintiff’s property. The damage was caused by pressurized smoke which the defendant repeatedly failed to adequately contain. The complaint alleged, and compensatory damages were awarded for, property damage. But, for the purpose of punitive damages, the court identified the state’s interest to be “deterring environmental pollution.” Based on this interest, the court allowed the jury to consider evidence regarding the defendant’s awareness that the smoke might cause cancer in humans and concluded that the punitive damage award reflected the reckless disregard for the health and safety of others exhibited by the defendant’s conduct, despite the fact that the plaintiff’s injury only involved damaged property.

*Action Marina* represents a gross abuse of the risk of harm doctrine; however, it is a logical extension of the concept. If courts are permitted to consider risks of harm to others that have no relationship to the plaintiff’s harm, then why even require a plaintiff to bring suit? Why require an award of compensatory damages to trigger a punitive damages award? Surely the law does not contemplate awarding punitive damages based, in part, on evidence that is not relevant to the specific injury for which standing was granted. While the defendant’s reckless disregard for the health and safety of others should be considered, it must be considered in relation to the harm suffered by the plaintiff.

Although, theoretically, the risk of harm to others does increase the reprehensibility of certain conduct, the tort system is not the appropriate method for punishing a defendant for risking harm to nonparties. Risk of harm to others is too speculative and is not linked to the plaintiff’s harm. Thus, it should not be considered when assessing the reprehensibility of a defendant’s conduct.

**B. Repeated Conduct**

In general, courts have used the concept of repeated conduct to describe a wide spectrum of evidence – testimony as to previous business practices, testimony regarding the number of times sparks from a defendant’s train
threatened the plaintiff’s crops,\textsuperscript{153} prior felony convictions,\textsuperscript{154} and even a consent order closing enforcement proceedings against an airline for eleven prior claims of racial discrimination.\textsuperscript{155}

Repeated conduct, as used by the courts, is overbroad. The concept encompasses three categories of conduct: “repeated acts” by the defendant against the plaintiff in the same course of conduct, “nonadjudicated repeated conduct” committed by the defendant outside the course of conduct at issue, and “adjudicated repeated conduct” committed by the defendant outside the course of conduct and for which the defendant was found liable/guilty.

1. Repeated Acts

Courts have consistently focused on the “frequency of the wrong” when upholding large punitive damage awards.\textsuperscript{156} However, many older cases focused on the frequency of the wrong, not against strangers to the litigation, but against the plaintiff.\textsuperscript{157} For example, in \textit{Gila Water Co. v. Gila Land & Cattle Co}, the court stated that cutting off the plaintiff’s water supply for a prolonged period was a “repeated diversion” of his water supply, and thereby warranted punitive damages.\textsuperscript{158} Later courts broadened the use of “repeated conduct” to embody any evidence of past conduct.\textsuperscript{159} But, there is a difference between conduct directed at the plaintiff during the course of conduct and conduct directed at others in the past. Courts should therefore distinguish repeated acts from repeated conduct for the purposes of determining reprehensibility.

The term “repeated acts” first appeared in \textit{State Farm}, which listed “the conduct involved repeated actions or was an isolated incident” as one of the five reprehensibility factors.\textsuperscript{160} Because \textit{State Farm} supported this factor with cases discussing repeated conduct,\textsuperscript{161} the case law tells us nothing about the meaning or appropriate use of repeated acts evidence.

\begin{itemize}
\item \textsuperscript{154} Ewing v. California, 538 U.S. 11, 17-20, 29 (2003).
\item \textsuperscript{157} \textit{See Guttell & Harel, supra note 156, at 1221.}
\item \textsuperscript{158} \textit{Gila Water Co. v. Gila Land & Cattle Co.}, 249 P. 751, 754-55 (Ariz. 1926).
\item \textsuperscript{159} \textit{See Pacific Mut. Life Ins. Co. v. Haslip}, 499 U.S. 1, 21-24 (1990) (approving the Supreme Court of Alabama’s punitive damage test that examines “the existence and frequency of similar past conduct”).
\item \textsuperscript{161} \textit{Id.} (citing BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 576-77 (1996)).
\end{itemize}
For our purposes, “repeated acts” means the defendant’s repeated actions which impacted the plaintiff during the same course of conduct. Under this construction, evidence of repeated acts would be appropriate for determining the reprehensibility of the defendant’s conduct. For example, in Philip Morris, evidence of every new tactic Philip Morris used – magazine ads, radio announcements, deceptive packaging, etc – which likely impacted Williams would be permissible for determining reprehensibility.

2. Nonadjudicated Repeated Conduct

Nonadjudicated repeated conduct embodies the Court’s existing, general notion of repeated conduct. The Federal Rules of Evidence suggest that evidence of repeated conduct should not be admissible to increase the overall reprehensibility of the defendant’s conduct.162 Fed. R. Evid. 404(b) provides that “evidence of other crimes, wrongs, or acts is not admissible” to show a defendant’s bad character, or to show that the defendant committed the conduct in question in accordance with that bad character.163 In other words, other acts evidence cannot be used to show that the defendant has a “propensity” to do the bad thing of which he is accused.164 However, such evidence is admissible for other purposes, “such as proof of motive, opportunity, intent, preparation, plan, knowledge, identity, or absence of mistake or accident.”165

The primary purpose of Rule 404(b) is to ensure that juries do not punish a defendant for conduct that is not on trial.166 Many decisions have stressed the importance of the rule, finding that other acts evidence tends to distract the jury from the main issue167 and lead to a decision based on the belief that the defendant is a bad person, rather than on the merits of the case at hand.168

162. See Gash, supra note 64, at 1194.
163. Fed. R. Evid. 404(b). The rule has been amended to clarify its applicability to civil as well as criminal cases. Fed. R. Evid. 404 advisory committee’s note.
168. U. S. v. Turquitt, 557 F.2d 464, 468 (5th Cir. 1977). According to the California Law Revision Commission:
Character evidence is of slight probative value and may be very prejudicial. It tends to distract the trier of fact from the main question of what actually happened on the particular occasion. It subtly permits the trier of fact to reward the good man and to punish the bad man because of their respective characters despite what the evidence in the case shows happened.

1964 CAL. LAW REVISION COMM’N REP. at 615.
Because reprehensibility is part of the punishment analysis, other acts evidence would not be admissible for the purpose of determining reprehensibility.

Nonadjudicated repeated conduct should be treated as other acts for the purposes of Rule 404(b). As such, evidence of nonadjudicated repeated conduct is only appropriate for purposes such as establishing motive, intent, knowledge, or absence of mistake when proving the elements of the cause of action. In addition, in cases where negligence alone is sufficient to warrant compensatory damages, some evidence of nonadjudicated repeated conduct may be appropriate for determining whether the defendant’s conduct rises to the level necessary to trigger punitive damages. But, such evidence is not appropriate for consideration when determining the level of punishment.

The Court has inadvertently alluded that nonadjudicated repeated conduct is admissible for a proper purpose but not to assess reprehensibility. In State Farm, for example, the Court held that the jury used evidence of nonadjudicated repeated conduct to impermissibly punish State Farm for its nationwide practices. The evidence at issue was introduced during the punitive damages phase of the trial, allegedly not for punishing the defendant, but for the proper purpose of demonstrating State Farm’s motive and intent. Although the Court did not explicitly address the proper purpose issue in its opinion, the fact that the plaintiffs only introduced this evidence during the punitive damages phase, and not during the trial phase, where evidence of motive and intent would be most important, showed that the plaintiff probably introduced the nonadjudicated repeated conduct evidence for the improper purpose of punishment. This alone would have constituted sufficient error to remand for a new trial.

Non-adjudicated repeated conduct evidence should only be used for truly proper purposes during both phases of the trial and not for increasing the

169. Generally, actual harm to others will not apply under the rule because it is the product of the same course of conduct in question.
170. Edward Imwinkelried argues that “other acts” evidence should be admissible based on the “doctrine of chances. For example, in a discrimination case, the plaintiff should be allowed to offer evidence of past discriminatory acts by the defendant to show that, in the plaintiff’s case, the defendant was “motivated by a discriminatory animus.” Or, in a case where a nanny is on trial for murdering her charge, the prosecution should be allowed to introduce evidence of even other children who died in a similar manner under the defendant’s care. Imwinkelried, supra note 165, at 420-21. Accepting the doctrine of chances does not offend this paper’s proposition because the doctrine would only allow evidence for a proper purpose, such as intent, motive, or to prove an element of the cause of action. My argument is that while other acts evidence should be permitted for its proper purpose, once the plaintiff proves the elements of her cause of action, other acts evidence is not relevant to determining reprehensibility, except when it is adjudicated conduct as discussed in Part IV.
172. Gash, supra note 63, at 1234-35.
173. Id. at 1234-35.
174. If the appeals court determined that permitting the evidence constituted a harmful, abuse of discretion.
punitive damage award. In order to be a legitimate use, this evidence should be necessary to establish the elements of a cause of action and/or the requisite mens rea for triggering punitive damages. Once the jury determines that the defendant is liable and the conduct was malicious or wanton and reckless, there is no further appropriate purpose for this type of evidence. In addition, the jury should be instructed throughout the trial as to the limited applicability of the evidence, and the judge should be sure to weigh the probative value of “other acts” evidence against the danger of “unfair prejudice, confusing the issues or misleading the jury.”

3. Adjudicated Repeated Conduct

Evidence of repeated conduct should only be considered for determining reprehensibility as part of a recidivism scheme. In BMW v. Gore, State Farm v. Campbell, and Philip Morris USA v. Williams, the Court defended the consideration of all types of other wrongs with references to recidivism. This leads to the conclusion that recidivism is the only nonparty other wrong that legitimately impacts reprehensibility.

However, the Court does not complete the recidivism analysis in any of these three cases. Recidivism should not be used as meaningless support for upholding a vague other wrongs doctrine. Instead, if the Court is serious about using recidivism as a justification for considering evidence of other wrongs when determining reprehensibility, it should provide civil defendants with procedural protections similar to those afforded criminal defendants. To accomplish this, the Court should adopt a legitimate civil recidivism framework that enhances punitive damage awards based on evidence of adjudicated repeated conduct.

175. See Jonathan Gross, Defending “Pattern and Practice” Evidence in Punitive Damages Cases, 61 Def. Couns. J. 403, 406 (1994) (“A watchful eye should be maintained to prevent the plaintiff from using the same [other acts] evidence for purposes other than that which the court has approved as relevant.”).
176. See Fed. R. Evid. 105.
177. Fed. R. Evid. 403 (“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.”).
179. Repeated acts, which I have argued are relevant to reprehensibility, are wrongs committed against the plaintiff during the same course of conduct.
III. RECIDIVISM

A. Recidivism Background

A recidivist is defined as “one who has been convicted of multiple criminal offenses, usually similar in nature.”180 All fifty states, as well as the federal government, currently have criminal recidivism statutes which provide for enhanced sentencing based on the defendant’s prior criminal history.181

The criminal justice system uses prevention and deterrence as the primary justifications for increasing a defendant’s sentence for recidivism.182 On one hand, recidivism statutes keep habitual criminals off the streets.183 In addition, the threat of increased punishment for repeated crimes acts as a deterrent to both the individual wrongdoer and others similarly situated.184

Recidivism is relevant to reprehensibility, not simply punishment. Theoretically, because the recidivist has been punished in the past for similar criminal behavior, she should have known better than to violate the law again.185 Because the recidivist should have known better, she is more culpable than a first-time offender186 and her crime is more reprehensible. The criminal system is not punishing the recidivist for all of her past crimes, but is increasing the punishment for the latest offence because the severity of the first punishment(s) failed to deter her from violating the law again.187 In Witte v. United States, the Supreme Court stated that “the enhanced punishment imposed for the later offense ‘is not to be viewed as . . . additional penalty for the earlier crimes,’ but instead as ‘a stiffened penalty for the latest crime, which is considered to be an aggravated offense because [it is] a repetitive one.’” (citation omitted).188 This is the logic the Court uses to support the idea that evidence of other wrongs can increase reprehensibility without increasing punishment. But, if the Court is going to use this to justify its holding, it must complete the comparison by adopting a civil recidivism framework.

184. See Logan, supra note 183, at 1619.
186. Id. at 246-47.
187. Id. at 246.
B. The Court’s Use of Recidivism and why it is Insufficient

In the punitive damages context, the Supreme Court has repeatedly used recidivism statutes to support the proposition that juries should consider harm to others and repeated conduct when assessing the reprehensibility of the civil defendant’s conduct. In BMW v Gore, for example, Justice Stevens stated that “[o]ur holdings that a recidivist may be punished more severely than a first offender recognize that repeated misconduct is more reprehensible than an individual instance of malfeasance.” However, Gore did not place any standards or restrictions on the type of misconduct that can be considered.

Although the Court’s general logic makes sense, it misses an important distinction between the repeated conduct used under recidivism statutes and the repeated conduct used under the Court’s current punitive damages framework. Criminal recidivism statutes increase punishment based on a criminal defendant’s prior convictions. In most cases, sentencing guidelines rarely allow for increased punishments based on a mere suspicion or allegation of misconduct.

Under the Federal Sentencing Guidelines, a court typically increases a defendant’s sentence for past criminal behavior only when there has been a prior sentence and/or conviction. The term prior sentence means “any sentence previously imposed upon adjudication of guilt.” Although the guidelines allow courts to consider criminal conduct not resulting in criminal conviction if reliable information indicates that the criminal history category under-represents the seriousness of the defendant’s criminal history, courts are expressly prohibited from considering a prior arrest record alone.

Oregon, the state where the Philip Morris suit originated, utilizes a similar recidivism system in its sentencing guidelines. Oregon’s sentencing chart bases a defendant’s criminal history on adult felony convictions, certain misdemeanor convictions, and juvenile adjudications. Unlike the federal

190. Gore, 517 U.S. at 577.
191. See Logan, supra note 183, at 1620.
192. See generally O’Neill, supra note 185, at 248-50 ("The use of arrests, without convictions, as a means of establishing a criminal History Category score appears to run against the grain of simple fairness," and that "prior contacts with the criminal justice system that only involve an arrest do not receive criminal history points . . . .").
198. Id.
Based on the theoretical justifications for recidivism statutes, it would not make sense under sentencing guidelines to enhance the punishment of a defendant who has never before been punished. If the defendant has not been punished for his conduct, then the judicial system does not know whether or not the punishment will successfully deter the defendant’s misbehavior. This may result in over-deterrence or may violate the constitutional principle that the “punishment must fit the crime.”

In *Gore*, the Court came close to acknowledging the principle that the repeated conduct must be adjudicated before it can justify an enhanced award. After stating “that a recidivist may be punished more severely than a first time offender,” Justice Stevens emphasized the significance of the fact that there was “no evidence that BMW had persisted in a course of conduct after it had been adjudicated unlawful on even one occasion, let alone repeated occasions.” Instead, BMW changed its policy in three states prior to the decision in *Gore*, and changed it nationwide only five days after the trial reached a verdict. BMW, like a good wrongdoer, was deterred by the formal punishment, just as the principles of recidivism contemplate.

Despite this glimmer of understanding in *Gore*, the Court still justifies the inclusion of all other wrongs – including actual and potential harm – with recidivism. Justice Stevens’s dissent illustrates the Court’s flawed understanding of recidivism statutes. He argues that if courts are permitted under recidivism statutes to increase the penalty for a present crime because of past misconduct, then surely courts should be permitted to enhance the penalty for misconduct that has never been punished and which injured multiple victims. Yet, the fact that the conduct has never been punished is precisely why the court should not be able to enhance the punishment based on recidivism. Recidivism statutes are justified on the grounds that the defendant’s repetitive crime is more culpable

---

199. See id. at §213-008-0002(b).
200. See generally O’Neill, supra note 185, at 246-47.
202. See id. at 575 n.24.
203. Id. at 577.
204. Id. at 579.
205. Id. at 579 n.31.

It is no answer to refer, as the majority does, to recidivism statutes. In that context, we have distinguished between taking prior crimes into account as an aggravating factor in penalizing the conduct before the court versus doing so to punish for the earlier crimes. But if enhancing a penalty for a present crime because of prior conduct that has already been punished is permissible, it is certainly proper to enhance a penalty because the conduct before the court, which has never been punished, injured multiple victims.

*Id.* (internal citations omitted).
because she was not deterred by the previous punishment. If there has been no previous punishment, then an enhancement based on a failure to learn the lesson is not justified. Recidivism statutes are intended to reflect levels of culpability not to punish defendants for nonadjudicated wrongs the court thinks they might have gotten away with in the past.

Similarly, the reprehensibility of harming multiple victims is punished through separate punishments for each person harmed, not through recidivism-based enhancement. Imposing punishments on a first-time offender for each victim/incident in addition to imposing recidivism-based enhancements should not be allowed because that would result in over-punishment.

The civil recidivism framework treats first time offenders who harm multiple people in the same manner as the state of Oregon would treat a similarly situated criminal offender. Oregon’s Sentencing Guidelines provide that an upward departure from the sentencing guidelines is permissible when “the offense involved multiple victims or incidents.” However, “this factor may not be cited when it is captured in a consecutive sentence.” Because a defendant can be sued separately by each victim, he should be treated the same as a criminal defendant who receives consecutive sentences for harming multiple victims.

In addition, the civil recidivism framework comports with due process. While considering actual harm to others in the reprehensibility analysis violates due process, recidivism enhancements do not because they provide notice to the defendant and because the underlying conduct has been formally adjudicated. This former adjudication gave the defendant an opportunity to raise defenses, as Due Process requires. In the criminal recidivism context, the Court held that “[o]ther than the fact of a prior conviction, any fact that increases the penalty for a crime beyond the prescribed statutory maximum must be submitted to a jury, and proved beyond a reasonable doubt.” Importantly, prior convictions have already been proven beyond a reasonable doubt. Therefore, this exception for past convictions is not really an exception at all. Instead, the Court seems to be maintaining a standard that requires evidence to be proven true before it will be used to significantly increase a defendant’s sentence. Likewise, because reprehensibility directly impacts the level of punishment, and is often used to justify punitive damage awards beyond the reasonable relationship ratio (much like a sentence beyond the statutory maximum), only evidence of adjudicated repeated conduct is appropriate for consideration when making the reprehensibility determination.

207. O’Neill, supra note 185, at 246.
209. Id. at §213-008-0002(b)(G).
210. Id.
211. See supra Part II(A)(1).
Thus, the civil recidivism framework should only permit evidence of other
wrongs to increase a punitive damage award if it rises to the level of adjudicated
repeated conduct, meaning that the conduct resulted in civil liability and an
award of punitive damages or criminal sanctions. This would ensure that the
defendant had the opportunity to defend against the conduct, and it would ensure
that the conduct was proven to be malicious or wanton and reckless beyond at
least a preponderance of the evidence. 214

C. Civil Recidivism Enhancements - Extraterritoriality and Dissimilar Conduct

In addition to enumerating factors for determining reprehensibility, the State
Farm Court ruled on the admissibility of extraterritorial and dissimilar conduct.
At trial, the plaintiffs introduced a wide spectrum of evidence to prove that State
Farm had a nationwide pattern and practice of maximizing profits by capping
payouts on claims.215 This included evidence of: “[a]pplying an “appearance
allowance” for hail damage in Colorado . . . [t]he use of defense medical
examiners and medical records reviews in Arizona, Texas and Hawaii . . .
[h]i/low settlement agreements in California arbitrations . . . [a]justment
practices for earthquake claims in California by State Farm Fire & Casualty
Company, a separate corporate entity . . . [e]vidence regarding the prospective
cancellation of hurricane insurance in Florida . . . [a] $100,000,000.00 verdict in
Texas for an uninsured case . . . [and e]vidence of sixteen class actions in
Pennsylvania, Texas, Louisiana, California, Arkansas, Illinois, Arizona,
Michigan, Alabama, Georgia, Tennessee, and Washington.”216 The plaintiffs did
not produce any evidence indicating that these practices were illegal in the state
in which they occurred.217 The Supreme Court ultimately ruled that the evidence
should not have been admitted because the jury could not punish defendants for
lawful or unlawful conduct occurring in other states and because evidence of
repeated conduct must be similar to the conduct that injured the plaintiff.218 The
court reasoned that dissimilar, out-of-state evidence is impermissible because of
the lack of procedural safeguards associated with such evidence and the potential
violation of states’ sovereignty involved.219

Several scholars, including Wayne Logan and Rachel Van Cleave, reject the
Court’s decision that courts cannot punish for out-of-state conduct. They argue
that considering exterritorial evidence when awarding punitive damages is

214. My proposition requires the defendant to receive punitive damages or a comparable
criminal sanction in the previous adjudication rather than mere compensatory damages because
compensatory damages are not punitive in nature.
216. Solange E. Ritchie, State Farm v. Campbell: The Trend Towards Evidentiary and
217. Id.
218. Id.
219. Id. at 421.
comparable to considering criminal offenders’ out-of-state prior criminal history under recidivism statutes. And, if a criminal defendant’s prior out-of-state convictions can be considered for increased punishment, then a civil defendant’s conduct in another state should be considered as well.

Under the Court’s current conception of recidivism, which fails to differentiate the categories of other wrongs, this analogy is incomplete. In the criminal recidivism context the Court considers prior convictions, not just evidence of nonadjudicated repeated conduct or harm to others. However, if the Court adopts this paper’s prescription and enacts a civil recidivism framework, then Logan’s and Van Cleave’s criticisms would be more compelling; and, the Court’s reasons for disallowing the consideration of out-of-state conduct would warrant greater suspicion.

First, the State Farm Court seemed concerned with the lack of procedural safeguards for considering out of state evidence in the civil context. But, unlike the current approach, the civil recidivism framework provides concrete, procedural protections. This ensures that punitive damage awards are administered in a precise and systematic manner.

Second, the Court expressed concern that considering out-of-state conduct violates federalism. In State Farm, the Court held that “[n]or, as a general rule, does a State have a legitimate concern in imposing punitive damages to punish a defendant for unlawful acts committed outside of the State’s jurisdiction.” However, recidivism enhancements do not punish the defendant, but rather, reflect the repeat offender’s heightened culpability. If anything, recidivism enhancements are meant to work with, rather than supersede the other states’ laws.

Under the civil recidivism framework, juries, like sentencing judges, should be permitted to consider judgments, sentences, or fines entered against the defendant for similar conduct in other states. First, the forum court would conduct an “internal” review of the elements of the offense to determine if it would: (1) be cognizable in the forum state and (2) trigger punitive damages or criminal sanctions under the laws of the forum state. If so, then the jury would be permitted to hear evidence relating to the adjudicated conduct.

In addition to limiting the use of extraterritorial conduct, the State Farm Court held that repeated conduct must replicate the conduct at issue. The Court’s ruling may only apply to the category of evidence disallowed in State

220. Logan, supra note 183, at 1611-12; Van Cleave, supra note 65, at 270-71.
221. Logan, supra note 183, at 1611-12; Van Cleave, supra note 65, at 270-72.
224. Id. at 421.
225. See Logan, supra note 183, at 1628.
226. See generally id. at 1621.
227. State Farm, 538 U.S. at 424.
Farm – dissimilar nonadjudicated repeated conduct, which under the civil recidivism framework would not be relevant to reprehensibility anyway.

As discussed in Part II(B)(2), Federal Rule of Evidence 404(b) prohibits evidence of “other acts,” (i.e. nonadjudicated repeated conduct) unless it is relevant for a proper purpose. It is hard to see how past, dissimilar, nonadjudicated, out-of state conduct can be relevant to a proper purpose such as motive, intent, plan, or knowledge. In State Farm, the plaintiff offered evidence of lawful, out-of-state misdeeds for the purpose of proving motive. In rejecting that evidence as dissimilar, the Court seemed to adopt a “nexus to the plaintiff” test for determining the relevance of nonadjudicated conduct. This test holds that, if the out-of-state conduct is dissimilar and has no nexus to the plaintiff, it is not relevant to a proper purpose. In fact, such dissimilar evidence is probably offered for an improper purpose, such as increasing the punishment award.

Evidence of adjudicated repeated conduct, on the other hand, is not introduced for a purpose such as establishing motive or intent, nor is it introduced to punish the defendant. Rather, this evidence is introduced to show that the defendant is more culpable because he refuses to alter his bad behavior.

However, even adjudicated repeated conduct under the civil recidivism framework should replicate, to some degree, the conduct in question. The fact that a defendant is held liable for an unrelated tort does not indicate that he failed to learn a lesson. For example, if a company is hit with punitive damages in 1995 for violating an air pollution regulation and then violates an employment discrimination law in 2000, it is unlikely that the defendant failed to learn his lesson from the first award. Thus, the State Farm similarity requirement may be appropriate under the civil recidivism framework as well.

IV: PROBLEMS WITH MY PRESCRIPTION

The primary problem with my general prescription is the risk of under-deterring corporate defendants. Under my approach, each person harmed by the defendant’s conduct should be entitled to the same slice of punitive damages. If ten people are injured, bring suit, and are awarded punitive damages, the defendant will receive his full punishment. However, if any or all of the potential plaintiffs fail to bring suit or are unsuccessful, the defendant will escape without receiving the full punishment. This problem could be addressed through

228. Fed. R. Evid. 404(b).
230. Id.
231. Id. at 423-24.
232. Logically, this nexus to the plaintiff standard should apply for all nonadjudicated repeated conduct evidence, not just evidence from other states.
233. In fact, it is this author’s contention that the same should be true of criminal recidivism statutes.
modifications of the plan, such as adding mandatory class actions.\textsuperscript{234} In addition, the government could address this by increasing criminal investigation and sanction of corporate misconduct.

The risk of under-deterrence extends to my proposed reprehensibility analysis, as well. Excluding harm to others and nonadjudicated repeated conduct from the reprehensibility analysis may result in the under-assessment of reprehensibility. If reprehensibility is under-assessed, even if every potential plaintiff brings suit, the overall punishment will be less than what the defendant’s conduct deserved.

The civil recidivism framework also presents several problems. First, requiring a formal adjudication of liability may limit repeated conduct evidence too much. In addition, it may encourage settlement (which may or may not be desirable). The civil recidivism framework, while providing some procedural safeguards, will still be susceptible to the caprice of the jury. Once the evidence of recidivism is presented, the jury has total discretion to determine how much the reprehensibility level should be affected. This would fail to remedy the uncertainty and inconsistency of the current system. But, society may want the tort system to leave this determination to the sound discretion of the jury. If not, enhancement guidelines, much like criminal sentencing guidelines, could be crafted to assist the jury in deciding the appropriate amount of punishment.\textsuperscript{235}

In addition, it is unclear how the civil recidivism framework would relate to the reasonable relationship guidepost, which provides that punitive damages must be “reasonably related” to the amount of compensatory damages.\textsuperscript{236} My contention is that because the Court has not addressed repeated conduct within the civil recidivism framework, any award including a recidivism enhancement should not be limited by the single-digit multiplier principle. For example, if without consideration of adjudicated prior conduct, a defendant’s tort warrants a punitive damage award nine times greater than the compensatory damages, then the Court should permit the additional recidivism enhancement to push the

\textsuperscript{234} See Cabraser, supra note 131, at 1734-35; Sharkey, supra note 112, at 410-14. But see Nelson v. Wal-Mart Stores, Inc., 245 F.R.D. 358, 377-78 (E.D. Ark. 2007) (denying certification of a class on the grounds that “unless each alleged class member has actually suffered harm from the pattern of illegal acts . . . [the defendant] will be over-deterred by a classwide award of punitive damages. Individualized determinations are necessary to fully realize the extent of harm caused by [the defendant’s] conduct and properly assess the need for punishment and deterrence”).

\textsuperscript{235} For example, a simple guideline might instruct the jury to double the award of punitive damages for every similar, adjudicated prior conduct. This may seem like impermissibly punishing the defendant for repeated conduct. However, it is my contention that because reprehensibility is used to calculate the punishment award, there is no real distinction between punishment and reprehensibility.

\textsuperscript{236} See BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 580 (1996) (“[E]xemplary damages must bear a ‘reasonable relationship’ to compensatory damages . . . .”); State Farm Mut. Auto. Ins. Co. v. Campbell, 538 U.S. 408, 425 (2003) (“[F]ew awards exceeding a single-digit ratio between punitive and compensatory damages . . . will satisfy due process. In Haslip, in upholding a punitive damages award, we concluded that an award of more than four times the amount of compensatory damages might be close to the line of constitutional impropriety.” (internal citations omitted)).
award into a double-digit multiplier range. Otherwise, the civil recidivism framework will be too weak to be effective. Since my plan removes most other wrongs evidence from the reprehensibility analysis, the recidivism framework must have enough weight to deter and punish.

CONCLUSION

While the Supreme Court may have intended the *Philip Morris* decision to clarify when and how evidence of other wrongs can be used to assess punitive damages, it has only confused the issue. The Court needs to determine whether or not it intends for all other wrong evidence to be assessed in the same way.

Restricting the reprehensibility analysis will likely result in lower punitive damage awards. On one hand, it is important for the judicial system to adequately punish wrongdoers for their misdeeds. Yet, defendants deserve procedural and substantive protections to keep the process fair and the punishment equitable. My five categories of other wrongs and civil recidivism framework provide one method the Court could use to clarify its other wrongs doctrine and buttress the punitive damage determination with procedural and substantive safeguards.

APPENDIX:

INSTRUCTING THE JURY ON REPREHENSIBILITY

My model jury instruction attempts to incorporate everything discussed in parts II and III. The instruction does three things. First, it divides the punitive damage determination into two steps: (1) determining if punitive damages apply and (2) determining the actual amount of punitive damages. Each step contains specific factors for consideration. Second, the instruction places each of the five categories of other wrongs into the corresponding step. Third, it adds a recidivism instruction to the reprehensibility analysis.

A. Model Jury Instruction

In addition to awarding damages to compensate the plaintiff, you may, but you are not required to, award punitive damages. The purpose of punitive damages it to punish the defendant for (wanton and reckless, or malicious) acts and thereby to discourage the defendant and others from engaging in similar conduct in the future. You may award punitive damages if you find that the defendant’s conduct was malicious [meaning that the defendant intended to injure the plaintiff] or wanton and reckless [meaning that the defendant acted with a reckless and outrageous indifference to a highly unreasonable risk of harm and acted
with a conscious indifference to the health, safety, and welfare of others.]

Step One: In determining the **scope** of defendant’s conduct, and if the defendant’s conduct towards the plaintiff was malicious or wanton and reckless, you may consider:

- The duration of the conduct, whether others were *actually harmed* from the defendant’s conduct and whether the defendant’s conduct was likely to continue
- Whether the defendant had a *practice or pattern* of engaging in and profiting from wrongful conduct similar to that which injured the plaintiff
- The seriousness of the *risk of harm* to the public arising from the Defendant’s conduct

Step Two: When awarding an **amount** of punitive damages you may consider:

- The reprehensibility of defendant’s conduct towards the plaintiff
- That the punitive damages must bear a reasonable relationship to the actual and/or potential harm that defendant’s conduct caused to the plaintiff
- Whether or not criminal sanctions have been imposed on the defendant for his/her/its conduct against the plaintiff.

When determining the reprehensibility of the defendant’s conduct in step two, you may consider:

- The likelihood at the time that serious harm to the plaintiff would arise from defendant’s conduct and the degree of the defendant’s awareness of that likelihood
- Whether the harm caused to the plaintiff was physical as opposed to economic
- Whether the plaintiff was financially vulnerable
- Whether the defendant’s conduct towards the plaintiff involved *repeated actions* or was an isolated incident
- Whether the defendant is a *recidivist*. A recidivist is defined as “one who has been convicted of multiple criminal offenses, usually similar in nature. A Recidivist is more culpable than a first-time offender. When determining if the defendant is a recidivist, you may only consider similar conduct for which
  - a jury held the defendant liable and awarded punitive damages, or
  - the defendant was adjudicated guilty and received civil or criminal sanctions.

If you determine that the defendant is a recidivist, you may increase the level of reprehensibility in proportion to the
number of instances and the seriousness of the defendant’s prior misconduct.

My model jury instruction is susceptible to criticism. Although the instruction addresses each category of other wrongs in the appropriate step of the analysis, there is still a concern that juries will consider the evidence for an improper purpose. Step one, where the jury determines if punitive damages are triggered and can consider all types of other wrongs, presents the greatest risk of this occurring. It may be that Step one of the instruction is not needed. Instead, that evidence should be used to prove mens rea during phase one of the trial. Step one would only be needed in situations where “intent” or “wanton and reckless” were not elements of the tort, and so, during phase two, the jury still had to determine if punitive damages were warranted. In addition, the trial judge should continue to play an important role in limiting irrelevant evidence and ensuring that evidence is only considered for its proper purpose.
FRANKFURTER’S GLOSS THEORY, SEPARATION OF POWERS, AND FOREIGN INVESTMENT

Catherine H. Gibson

In February and March 2006, members of Congress won a rare victory in foreign relations—they successfully opposed the President. Months earlier, executive officials had approved the acquisition of U.S. seaport management companies by Dubai Ports World, a United Arab Emirates-based corporation. Despite this executive approval, Congress thwarted the transaction by exerting political pressure based on national security concerns. Though some have criticized Congress’s actions, the Dubai Ports World controversy was actually a victory for Congress in the debate over separation of powers in foreign relations. In other foreign relations matters, such as international agreements, war, and surveillance, the President has gradually taken on more and more power. Because the Constitution does not comprehensively define the foreign relations powers of the legislature and the executive, the historical relationship between the branches, what Supreme Court Justice Felix Frankfurter called the gloss on executive power, often proves decisive to permanently shift power to the executive.

This article analyzes the historical relationship between Congress and the President in foreign investment and seeks to determine whether Congress can preserve its role in the area under current law. By examining the operation of Frankfurter’s gloss theory in international agreements, war, and surveillance, this paper establishes a framework of factors that weigh for and against continued congressional participation in foreign relations decisions. This framework is applied to the 2007 amendments to foreign investment legislation to assess the standing of separation of powers and to make policy recommendations.

Table of Contents

INTRODUCTION ................................................................................................... 105
I. SEPARATION OF POWERS IN FOREIGN INVESTMENT ................................. 107
   A. The 1930s to 1960s: Trade agreements and national security concerns ....................... 108
   B. 1970s: Executive Order 11858 and IEEPA ........................................... 108

* Law Clerk to the Honorable Samuel G. Wilson, U.S. District Court for the Western District of Virginia; J.D. and L.L.M., 2008, Duke University School of Law; B.A., 2004, Wake Forest University. For helpful comments, I thank Scott L. Silliman and the staff of the Northern Kentucky Law Review.
C. 1980s: Exon-Florio ................................................................. 110
D. 1990s: The Byrd Amendment .................................................... 112
E. 2000s: The Foreign Investment and National Security Act .... 113

II. SEPARATION OF POWERS IN OTHER FOREIGN RELATIONS MATTERS ... 116
A. International agreements ........................................................ 116
   1. Formation of international agreements .................................. 117
      a. Presidential assertions of power ..................................... 117
      b. Judicial validation ...................................................... 118
      c. Congressional reassertions of power ............................ 118
   2. Alternative forms of international agreements ..................... 120
      a. Presidential assertions of power .................................. 120
      b. Judicial validation ...................................................... 122
      c. Congressional reassertions of power ............................ 122
   3. Critical analysis .............................................................. 123
B. War .......................................................................................... 124
   1. Presidential assertions of power ..................................... 124
   2. Judicial validation ........................................................... 125
   3. Congressional reassertions of power .............................. 127
   4. Critical analysis .............................................................. 128
C. Surveillance ............................................................................. 128
   1. Presidential assertions of power ..................................... 129
   2. Judicial validation ........................................................... 129
   3. Congressional reassertions of power .............................. 131
   4. Critical analysis .............................................................. 133

III. HISTORICAL FRAMEWORK FOR SEPARATION OF POWERS ........ 134
A. Threshold consideration: constitutional language .................. 134
B. Factors for shifting power to the executive ............................ 135
C. Factors against shifting power to the executive ..................... 136

IV. APPLYING THE FRAMEWORK TO FOREIGN INVESTMENT .......... 137
A. Threshold consideration: constitutional language ................. 138
B. Review of current legislation ................................................. 138
   1. Congressional inaction ................................................. 138
   2. Judicial approval .......................................................... 139
   3. Congressional effectiveness .......................................... 140
C. Policy recommendations ....................................................... 140
   1. Individual comments .................................................... 140
   2. Legislative participation ............................................... 141
   3. Mandatory language .................................................... 141
   4. Media scrutiny ............................................................. 141

CONCLUSION .................................................................................. 142
INTRODUCTION

In February and March 2006, members of Congress won a rare victory in foreign relations—they successfully opposed the President. In mid-February, congressional leaders learned of a pending corporate transaction that would transfer control of U.S. seaports in New York, Baltimore, and Philadelphia to Dubai Ports World, a United Arab Emirates-based corporation.1 Dubai Ports World had gained the executive’s approval for the transaction without legislative involvement,2 and shareholders brought the deal to Congress’s attention only in a final effort to prevent the acquisition.3 Though Congress lacked legal authority to block the Dubai Ports World deal, members of Congress created significant political pressure on the deal by expressing strong opposition based on national security concerns.4 Within three weeks, the parties abandoned the deal.5 To prevent repetition of this situation, Congress formally strengthened its role in the regulation of foreign investment in 2007 by amending existing laws to require greater legislative participation in approving foreign acquisition of U.S. national security assets, such as seaports. Though some have criticized Congress’s actions for discouraging foreign investment,6 the outcome of the Dubai Ports World controversy was a success for Congress in the struggle for power in foreign relations.

The separation of powers principle is a primary basis for the U.S. Constitution.7 As James Madison wrote in the Federalist Papers, the Framers “contriv[ed] the interior structure of the government as that its several constituent parts may, by their mutual relations, be the means of keeping each other in their proper places.”8 Though this principle functions relatively well in domestic affairs, it is less effective in foreign relations, where clear differences between executive and legislative functions are difficult to define.9 For example,
though Congress declares war and the President acts as Commander in Chief, the Constitution does not indicate when the declaration powers end and the Commander in Chief powers begin. Indeed, the Constitution does not even mention many essential foreign relations powers, such as the power to terminate treaties or to negotiate peace agreements.

Without a clear constitutional framework, courts and scholars have come to rely on other sources to determine the appropriate relationship between the executive and legislative branches in foreign relations. One such source is the historical relationship between the two branches, called the gloss of history on executive power. As stated by Justice Felix Frankfurter in his concurrence to *Youngstown Sheet & Tube Co. v. Sawyer*:

> [A] systematic, unbroken, executive practice, long pursued to the knowledge of the Congress and never before questioned, engaged in by Presidents who have also sworn to uphold the Constitution, making as it were such exercise of power part of the structure of our government, may be treated as a gloss on ‘executive Power’ vested in the President.\(^\text{10}\)

This theory is effective because, according to Louis Henkin, “[t]he absence of a comprehensive, ‘natural’ division of political authority in foreign affairs which would indisputably exclude the other political branch has been a strong invitation to compete for power, and to claim at least a concurrent authority.”\(^\text{11}\) Under Frankfurter’s gloss theory, because Congress and the President compete for foreign relations powers, congressional acquiescence to the exercise of a certain power by the President is evidence that power belongs to the executive.

Though the Framers generally favored a balance of powers among branches of government, in foreign relations Frankfurter’s gloss theory has caused a gradual shift of power to the executive.\(^\text{12}\) Because the executive enjoys the advantages of expediency, unity, and secrecy, Presidents have gained foreign affairs powers particularly in national emergencies, when these advantages are most important. Specifically, this shift may be seen in executive control of international agreements, war, and surveillance. With respect to foreign investment in U.S. national security assets, however, the executive has declined to exercise such broad powers, even in times of national emergency. Instead, the President continues to share foreign investment powers with Congress, and the relationship between the two branches has evolved by legislative amendments as new problems have arisen. Therefore, though Congress’s frustration of the tools such as the political question doctrine and the act of state doctrine to avoid deciding the merits of controversial foreign affairs cases. *Id.* at 137-48.

\(^\text{10}\text{343 U.S. 579, 610-11 (1952).}\)

\(^\text{11}\text{HENKIN, supra note 9, at 93.}\)

\(^\text{12}\text{LOUIS FISHER, CONSTITUTIONAL CONFLICTS BETWEEN CONGRESS AND THE PRESIDENT 287 (5th ed. 2007) [hereinafter CONSTITUTIONAL CONFLICTS]. As Fisher writes, “[t]he general drift of authority and responsibility toward the President over the past two centuries is unmistakable.” Id.}\)
Dubai Ports World deal over executive approval may have injured U.S. trade interests, its success is a victory for Congress in the struggle for power over U.S. foreign relations.

The 2007 amendments to U.S. foreign investment laws—passed partially in response to the Dubai Ports controversy—are, however, a marked change to previous policy. These amendments expand executive discretion to determine whether foreign investments threaten national security but delegate authority directly to an executive committee, rather than to the President. This article analyzes the history of separation of powers in foreign investment and seeks to determine whether current laws will permit Congress to maintain its role in foreign investment under Frankfurter’s gloss theory on executive power.

The article proceeds as follows: Part I describes the development of U.S. law regulating foreign investment in national-security assets, particularly noting legislative responses to changes in the investment environment. Part II discusses the history of separation of powers in international agreements, war, and surveillance, focusing on executive assertions of power in each area and the relative success of congressional attempts to regain partial control. Part III presents a framework of factors for and against the drift of foreign relations power to the President based on the operation of Frankfurter’s gloss theory in international agreements, war, and surveillance. Finally, Part IV applies this framework to the current foreign investment regime, noting the aspects of these laws that secure and threaten the existing separation of powers and offering policy recommendations to ensure a continued congressional role in foreign investment.

I. SEPARATION OF POWERS IN FOREIGN INVESTMENT

Like other foreign relations issues, foreign investment in national security assets has long been a point of concern to both Congress and the President. Unlike other foreign relations powers, however, the executive and legislative branches have generally shared control of foreign investment, reaching new compromises as new needs arise. The federal government’s control over foreign investment began as a relatively informal monitoring mechanism. As foreign investment became more common and complex, Congress and the President increased the federal government’s power by requiring parties to report certain transactions and by allowing the government to block or modify transactions deemed national security threats.

At present, by agreement of both branches, the government body principally charged with weighing foreign investment considerations is a group of executive branch officials called the Committee on Foreign Investment in the United States (CFIUS or the Committee). Under the current statutory framework, CFIUS may investigate, analyze, and block foreign investment. Regardless of its action on a particular transaction, the Committee must report its activities and decisions to Congress. The committee’s powers, however, were not always so broad. Since
the Committee’s creation in 1975, Congress has amended the laws governing its powers to accommodate changes in the investment environment.

A. The 1930s to 1960s: Trade agreements and national security concerns

From the 1930s to the 1960s, Congress granted the President broad powers to conclude trade agreements with foreign governments.13 As trade relations began affecting national security, Congress adjusted these powers accordingly. First, in 1954, Congress passed a resolution prohibiting the President from reducing duties on products if increased import levels would threaten “domestic production needed for projected national defense requirements.”14 Congress also created an investigative mechanism to determine the national security effects of trade agreements.15 Finally, Congress permitted the President to take any action necessary to address threats to national security in trade agreements.16 Though these early statutes affected only agreements between the U.S. government and the governments of other nations, they provided a framework for the future regulation of private transactions by CFIUS.

B. 1970s: Executive Order 11858 and IEEPA

In 1974, Congress passed the Foreign Investment Study Act, which required the executive to review foreign investment in the United States.17 After this review revealed that the federal government lacked a comprehensive mechanism to monitor foreign investment, the President issued an executive order in 1975 creating CFIUS and charging that Committee with monitoring foreign investment and consulting with the private entities concerned in such transactions.18 The Committee, however, lacked authority to block transactions.19

13. For example, the Smoot Hawley Tariff Act of 1930 was amended in the 1950s to permit the President to intervene in private trade agreements that threatened national security. For a longer discussion of these acts, see Deborah M. Mostaghel, *Dubai Ports World Under Exon-Florio: A Threat to National Security or a Tempest in a Seaport*, 70 AM. L. REV. 583, 586-87 (2006).
16. Id. For example, in response to a perceived threat to the U.S. machine tool industry caused by imports, President Reagan concluded Voluntary Restraint Agreements with two of the primary importers of machine tools under which these nations agreed to reduce their imports to the United States. Edward E. Groves, *A Brief History of the 1988 National Security Amendments*, 20 LAW & POL’Y INT’L BUS. 589, 594-97 (1988). These agreements were deemed necessary for national security purposes since machine tools are necessary for the production of weapons to supply the U.S. military. Id.
18. Exec. Order No. 11,858, 3 C.F.R. 990 (1975), reprinted in 15 U.S.C. § 78(b) app. at 163-65 (2000). Specifically, the order required the Committee to analyze “trends and significant developments” in the United States in foreign investment, advise on “prospective major foreign government investments,” review investments which might “have major implications for United
As the Executive Order provided, CFIUS “shall have primary continuing responsibility within the Executive Branch for monitoring the impact of foreign investment in the United States, both direct and portfolio, and for coordinating the implementation of United States policy on such investment.” In addition, though “[a]ll departments and agencies [were] directed to provide . . . information and assistance as may be requested by the Committee,” the Committee’s own reporting requirements were discretionary.

Two years later, Congress codified and expanded the executive’s authority over foreign investment in the International Economic Emergency Powers Act (IEEPA). IEEPA empowered the President to investigate or block transactions in foreign currency or transactions involving property owned by a foreign country or a foreign national. Under IEEPA, however, the President could exercise these powers only in a declared national emergency. Because Presidents proved unwilling to declare a national emergency to prevent individual transactions, Congress later relaxed this requirement in the Exon-Florio provision of the Omnibus Trade and Competitiveness Act.

States national interests, and consider proposals for new foreign investment legislation or regulations. Id. at 991.

19. Id.
20. Id.
21. Id.
23. Id. § 1702. The statute provides,

(1) At the time [of a declared national emergency], the President may . . .

(A) investigate, regulate, or prohibit--

(i) any transactions in foreign exchange,

(ii) transfers of credit or payments between, by, through, or to any banking institution, to the extent that such transfers or payments involve any interest of any foreign country or a national thereof,

(iii) the importing or exporting of currency or securities,

by any person, or with respect to any property, subject to the jurisdiction of the United States;

(B) investigate, block during the pendency of an investigation, regulate, direct and compel, nullify, void, prevent or prohibit, any acquisition, holding, withholding, use, transfer, withdrawal, transportation, importation or exportation of, or dealing in, or exercising any right, power, or privilege with respect to, or transactions involving, any property in which any foreign country or a national thereof has any interest by any person, or with respect to any property, subject to the jurisdiction of the United States . . .

24. Id. § 1701(a). The President could declare a national emergency in response to an “unusual and extraordinary threat . . . to the national security, foreign policy or economy of the United States.” Id.
C. 1980s: Exon-Florio

Exon-Florio, enacted in 1988, expands executive authority over foreign investment by lowering the threshold requirement for executive action and by granting the President wide discretion in determining the existence of “national security” threats that trigger CFIUS’s powers. Exon-Florio was an indirect response to the proposed merger of Japan’s Fujitsu, Ltd. and the Fairchild Semiconductor Corporation of the United States, both producers of semiconductors. Because the merger would have made U.S. military technology dependent on Japanese semiconductor production, executive officials viewed the Fujitsu-Fairchild deal as a national security threat. Under the 1974 Executive Order and IEEPA, however, neither the President nor Congress had legal authority to prevent the merger unless a national emergency was declared. The deal was eventually blocked through informal political opposition, and Congress amended U.S. regulation of foreign investment to prevent the repetition of a similar situation.

In response to the Fujitsu-Fairchild controversy, Exon-Florio created a mechanism that allows the President to block transactions that threaten national security even without declaring a national emergency. Under Exon-Florio, the President can suspend or prohibit any transaction with foreign entities based on findings that:

1. there is credible evidence that . . . the foreign interest exercising control might take action that threatens to impair the national security;
2. provisions of law, other than this section and [IEEPA], do not in the President’s judgment provide adequate and appropriate authority for the President to protect the national security in the matter before the President.

Thus, Exon-Florio expands the President’s authority over foreign investment by lowering the threshold findings required for executive action under the statute. Rather than requiring a declared national emergency as in IEEPA, Exon-Florio only requires credible evidence of a national security threat.

In addition, because Exon-Florio does not specifically define “national security,” the statute allows the President significant discretion in determining
when to act. In finding a national security threat under Exon-Florio, the President must consider factors such as “domestic production needed for projected national defense requirements,” “the capability and capacity of domestic industries to meet national defense requirements,” and “the control of domestic industries and commercial activity by foreign citizens as it affects the capability and capacity of the United States to meet the requirements of national security.” The statute, however, does not mandate the weight given to the presence or absence of each factor, therefore permitting the President to exercise wide discretion.

The President’s “national security” determinations are subject to only limited review. Though Congress receives prompt reports of any action taken under Exon-Florio, the statute exempts the President’s national security determinations from judicial review and exempts the information received in making these determinations from disclosure under the Freedom of Information Act (FOIA). Thus, neither the courts nor the public may review the President’s national security determinations.

Significantly, however, the President’s authority under Exon-Florio is permissive. The executive has discretion to determine whether a particular transaction warrants investigation or interference, but is not required to act under any circumstances. In addition, Exon-Florio delegates this authority directly to the President, who then delegated it to CFIUS. Though some scholars feared that Exon-Florio would lead to significant government interference in the economy, CFIUS investigated only seventeen of nearly 1,300 foreign-investment transactions voluntarily reported to the executive pursuant to the statute between 1988 and 1999. Of these seventeen investigations, only one transaction was blocked.

33. Id.
34. See id.
35. Id.
36. Id. Exon-Florio also provides a strict timeline for the President’s actions: Investigations must be initiated within 30 days after the President receives notice of the proposed transaction, investigations must be completed within 45 days, and any action by the President to block the transaction must take place within 15 days of the completion of the investigation. Id.
38. UNITED STATES GENERAL ACCOUNTING OFFICE, DEFENSE TRADE: IDENTIFYING FOREIGN ACQUISITIONS AFFECTING NATIONAL SECURITY CAN BE IMPROVED 7-8 (2000).
39. In that transaction, President George H. W. Bush ordered the China National Aero-Technology Import and Export Corporation to divest its interests in MAMCO Manufacturing, Inc.,
D. 1990s: The Byrd Amendment

In 1993, Congress amended the Exon-Florio framework in a provision of the Defense Authorization Act commonly called the Byrd Amendment.40 This amendment was primarily a response to a controversial bid for the aerospace and defense operations of the American-owned LTV Corporation by Thomson CSF Inc., a corporation majority-owned by the French government.41 Though foreign corporations had made similar acquisitions of U.S. entities in the past, the Thomson bid was particularly controversial because LTV was involved in the development of top-secret military defense technology.42 As in the Fujitsu-Fairchild deal, Thomson CSF withdrew its bid for LTV after intense media scrutiny and congressional consternation.43

In response to the LTV-Thomson bid, the Byrd Amendment made several changes to the Exon-Florio framework. First, the Byrd Amendment required the executive to investigate transactions involving a foreign entity that is “controlled by or acting on behalf of a foreign government.”44 Thus, under the Byrd Amendment, executive review of the LTV-Thomson transaction would have been mandatory due to the involvement of the French government.

Because the LTV-Thomson transaction would have allegedly compromised U.S. military technology secrets, the Byrd Amendment also added new safeguards for technology. The amendment introduced two new factors for consideration in national security determinations: the effect of the proposed transaction on both U.S. “technological leadership in areas affecting United States National Security” and “sales of military goods, equipment, or technology to any country,” particularly to countries posing concerns regarding missile or chemical weapons proliferation, or a country that supports terrorism.45 In addition, the amendment requested that the Director of the Office of Science and

an American manufacturer of commercial aircraft components. Christopher R. Fenton, Note, U.S. Policy Towards Foreign Direct Investment Post-September 11: Exon-Florio in the Age of Transnational Security, 41 Colum. J. Transnat’l L. 195, 210 (2002). Because the Chinese corporation involved in this transaction was owned by China’s Ministry of Aerospace Technology, a close affiliate of the People’s Liberation Army, the President deemed this transaction a security risk. Id.

42. George Graham, French missile bid stirs row: George Graham on U.S. anger at Thomson-CSF’s move to buy missiles division of U.S. defense contractor, FINANCIAL TIMES, May 22, 1992, at 5. Allegations that France had sold arms to Iraq before the first Gulf War also added to the controversy. Key senators oppose restructured LTV-Thomson deal, 163 AEROSPACE DAILY, July 7, 1992, at 32.
44. 50 U.S.C. app. § 2170.
45. Id.
Technology Policy and the Assistant to the President for National Security be added as members of CFIUS.\(^{46}\)

Finally, the Byrd Amendment strengthened reporting requirements, thereby reasserting congressional power in foreign investment. Under the statute, the President must provide “a detailed explanation” of the national security factor analysis and must issue reports when he chooses not to act on a particular transaction, rather than only when he acts.\(^{47}\)

E. 2000s: The Foreign Investment and National Security Act

In 2007, Congress amended the Exon-Florio framework again by enacting the Foreign Investment and National Security Act (FINSA).\(^{48}\) FINSA is a response to the attacks of September 11, 2001, the Dubai Ports World controversy, and a failed bid by the state-owned China National Offshore Oil Corporation (CNOOC) for the American energy producer Unocal Corporation.\(^{49}\) To address the CNOOC bid, the President sought to simply follow the CFIUS review process, but Congress passed a special resolution to frustrate this particular acquisition.\(^{50}\) In response, CNOOC withdrew its bid for Unocal, even before the CFIUS process was completed.\(^{51}\)

FINSA responds to 9/11, the Dubai Ports World controversy, and the CNOOC bid by providing a more rigid review process and by adding factors to the President’s national security determinations. Under FINSA, the Committee’s involvement begins when the executive receives notice of a transaction “by or with any foreign person which could result in foreign control of any person engaged in interstate commerce.”\(^{52}\) These notices typically include “a detailed description of the transaction and the parties, related government contracts, future plans for the U.S. entity, and foreign government involvement.”\(^{53}\)

\(^{46}\) Id.

\(^{47}\) Id. (emphasis added).


\(^{50}\) Id. at 163-64.

\(^{51}\) Id. at 164.

\(^{52}\) 50 U.S.C.A. app. § 2170(b) (West 2007). Though notification by the parties is optional, foreign acquirers are advised to make these notifications because CFIUS may initiate its own investigations and the Committee monitors news reports for transactions of interest. Victor Lewkow, Congress Tightens Exon-Florio “National Security” Reviews of Foreign Investment in the United States, 1647 PLI/Corp 463, 468-69 (2008). In addition, if the parties do not voluntarily notify the Committee of the transaction, CFIUS retains the authority to undo the transaction at any time, even after its conclusion. Id. at 469.

\(^{53}\) Lewkow, supra note 52, at 468.
FINSA establishes a three-tiered screening system of review, analysis, and investigation. First, CFIUS must review each reported transaction to determine its effects on national security. If the initial review indicates that a particular transaction threatens national security, FINSA’s second tier of screening, an analysis, is triggered. The analysis must be “expeditiously carr[ied] out” by the Director of National Intelligence (DNI). In completing an analysis, the DNI must engage the intelligence community and ensure that it provides CFIUS with any relevant information.

Finally, the third tier of screening, a national security investigation, is triggered when a transaction’s initial review reveals national security problems that have not been mitigated by prior adjustments, or if the transaction would result in foreign government control of a U.S. entity or foreign control of U.S. critical infrastructure. In an investigation, FINSA authorizes the committee to “take any necessary actions in connection with the transaction to protect the national security of the United States.”

As in Exon-Florio and the Byrd Amendment, “national security” determinations under FINSA are based on a list of factors, but are left largely to executive discretion. FINSA also adds several factors to the existing list,

55. Id. § 2170(b)(4).
56. Id. § 2170(b)(4)(A).
57. Id. § 2170(b)(4)(A)(4).
58. Id. § 2170(b)(2)(B)(i)(III).
59. Id. § 2170(b)(2)(A).
60. The complete list of factors provides:
[for the purposes of this section, the President or the President’s designee may, taking into account the requirements of national security, consider—
(1) domestic production needed for projected national defense requirements
(2) the capability and capacity of domestic industries to meet national defense requirements, including the availability of human resources, products, technology, materials, and other supplies and services,
(3) the control of domestic industries and commercial activity by foreign citizens as it affects the capability and capacity of the United States to meet the requirements of national security,
(4) the potential effects of the proposed or pending transaction on sales of military goods, equipment, or technology to any country—
(A) identified by the Secretary of State—
(i) . . . as a country that supports terrorism;
(ii) . . . as a country of concern regarding missile proliferation; or
(iii) . . . as a country of concern regarding the proliferation of chemical and biological weapons;
(B) identified by the Secretary of Defense as posing a potential regional military threat to the interests of the United States; or
(C) listed . . . on the “Nuclear Non-Proliferation-Special Country List” or any successor list;
(5) the potential effects of the proposed or pending transaction on United States international technological leadership in areas affecting United States national security;
including whether the foreign country implicated in a transaction is “a potential regional military threat” or whether that country has adhered to nonproliferation controls and cooperated with the United States on counterterrorism.\textsuperscript{61} In response to the CNOOC deal particularly, FINSA adds to national security determinations the “major energy assets”\textsuperscript{62} and “long-term projection of the United States requirements for sources of energy”\textsuperscript{63} as part of the domestic infrastructure considered valuable to preserving national security.

After completing its screening process, CFIUS must issue a report to congressional leaders and to the parties to the proposed transaction.\textsuperscript{64} If the President finds that a given transaction “threatens to impair the national security of the United States,” he “may take such actions for such time as [he] considers appropriate to suspend or prohibit [that] transaction.”\textsuperscript{65} Under current law, the full members of the Committee are the Secretaries of Treasury, Homeland

\begin{enumerate}
\item the potential national security-related effects on United States critical infrastructure, including major energy assets;
\item the potential national security-related effects on United States critical technologies;
\item whether the covered transaction is a foreign government-controlled transaction . . . ;
\item as appropriate . . . a review of the current assessment of—
  \begin{enumerate}
  \item the adherence of the subject country to nonproliferation control regimes, including treaties and multilateral supply guidelines, which shall draw on, but not be limited to, the annual report on “Adherence to and Compliance with Arms Control, Nonproliferation and Disarmament Agreements and Commitments” required by section 2593a of Title 22;
  \item the relationship of such country with the United States, specifically on its record on cooperating in counter-terrorism efforts, which shall draw on, but not be limited to, the report of the President to Congress under section 7120 of the Intelligence Reform and Terrorism Prevention Act of 2004; and
  \item the potential for transshipment or diversion of technologies with military applications, including an analysis of national export control laws and regulations;
  \end{enumerate}
\item the long-term projection of United States requirements for sources of energy and other critical resources and material; and
\item such other factors as the President or the Committee may determine to be appropriate, generally or in connection with a specific review or investigation.
\end{enumerate}

61. Id. § 2170(f)(4)(B).
62. Id. § 2170(f)(6).
63. Id. § 2170(f)(10).
64. Id. § 2170(b)(3). Reports must be issued to the Senate Majority and Minority Leaders; the Speaker and Minority of the House of Representatives; the members of the Senate from the state in which the principal place of business of the acquired U.S. person is located, and the member from the Congressional District in which such principal place of business is located; the chair and ranking member of the Committee on Banking, Housing and Urban Affairs of the Senate; and the Chair of the Committee on Financial Services of the House of Representatives. Id.
Security, Commerce, Defense, State, and Energy, and the Attorney General.\footnote{Id. § 2170(k)(2).} The DNI and the Secretary of Labor are non-voting, ex officio members of the committee.\footnote{Id. § 2170(k)(2).}

Therefore, on one hand, FINSA continues the policies of prior legislation by leaving national security determinations to executive discretion and by using reporting requirements to ensure congressional participation in foreign investment decisions. On the other hand, FINSA breaks from its predecessors by delegating power to CFIUS itself, rather than to the executive branch, and by permitting a single executive official, the DNI, to act alone in analyzing foreign investment proposals. The question remains, however, whether these provisions permit Congress to continue its role in foreign investment, or whether the President will gradually take control over this area through the operation of Frankfurter’s gloss theory, as has occurred in other areas of foreign relations law.

II. SEPARATION OF POWERS IN OTHER FOREIGN RELATIONS MATTERS

As demonstrated above, separation of powers between Congress and the President has been maintained in foreign investment. In other foreign relations matters, however, control is largely exercised by the President alone. Specifically, control of international agreements, war, and surveillance has largely shifted to the executive under Frankfurter’s gloss theory. This section analyzes changes in the relationship between Congress and the President in each of these three areas, and seeks to determine the factors that weigh for and against maintaining separation of powers in foreign relations.

A. International agreements

Under Article II of the Constitution, “[the President] shall have Power, by and with the Advice and Consent of the Senate, to make Treaties, provided two thirds of the Senators present concur.”\footnote{U.S. Const. art. II, § 2.} Though the placement of the treaty clause in Article II (governing executive powers) rather than in Article I (governing legislative powers) may indicate that the Framers contemplated a greater role for the executive in treaty-making, the clause nevertheless requires a role for the Senate in both advice and consent. In fact, the Constitution envisions strong Senate participation in treaty-making by requiring approval by more than a simple majority.

Despite this constitutional language, Frankfurter’s historical gloss on executive power has shifted authority over international agreements in favor of the President in two major ways. First, the President now negotiates international agreements largely on his own, without Senate advice. In addition,
the President now chooses whether to submit the product of his negotiations to the Senate as a treaty, requiring approval by a supermajority, or to Congress as an executive agreement, requiring only a simple majority. Courts have generally affirmed the President’s powers in these two areas, and Congress has been only partially successful in its attempts to reassert its power in these two respects.

1. Formation of international agreements

a. Presidential assertions of power

Originally, under Article II the “advice [of the Senate] . . . was to be given at every stage of diplomacy, from the framing of policy and instructions to the final bestowal of consent.”69 However, even during George Washington’s presidency this system proved unworkable: In 1789 Washington attempted to consult the Senate on a treaty with Southern Indians, but soon abandoned these efforts after consultation proved unwieldy.70 The Senate also proved unable to safeguard the confidential details of U.S. foreign relations when the terms of the 1795 Jay Treaty with Great Britain were leaked to a Washington, D.C. newspaper, though they were intended to be kept secret even after the treaty was concluded.71 Because treaty-making requires secrecy and dispatch,72 the Washington administration began seeking Senate consent only after treaties were drafted. Subsequent Presidents have also followed this approach. Thus, even the Senate now concedes that “the first phase of treatymaking, negotiation and conclusion, is widely considered the exclusive prerogative of the President.”73 Because the President often pressures the Senate to consent to treaties as they are submitted, exclusion from the drafting process particularly weakens the Senate.74

70. Id. at 1280-82.
71. Id. at 1282.
72. THE FEDERALIST NO. 64 (John Jay).
74. Id. at 15. Specifically, 
[t]he President may present a treaty as vital to good relations with a nation, relations that would be set back immeasurably if the treaty were defeated. Or he may present it as a package that has been so delicately negotiated that the slightest change in understanding by the Senate would unbalance the package and kill the treaty. Or he may present it so late in the congressional session, or so near some type of international deadline, that Senate consideration in depth is pictured as impeding the beginning of a new beneficial regime. 

Id. at 15-16.
b. Judicial validation

Courts have approved solo treaty negotiations by the executive, and have used this negotiation authority as a justification for granting the President power over the interpretation of treaties as well. As the Supreme Court stated in United States v. Curtiss-Wright Export Corp., “the President alone has the power to speak or listen as a representative of the nation. He makes treaties with the advice and consent of the Senate; but he alone negotiates.” 75 With respect to treaty interpretation, the Court has held that, “[w]hile courts interpret treaties for themselves, the meaning given them by the departments of government particularly charged with their negotiation and enforcement is given great weight.” 76 Thus, the Court has declined to enforce the specific language of the Constitution requiring Senate “advice” in the negotiation and conclusion of treaties. In fact, the Court has legitimated the President’s unilateral treaty-drafting by granting the views of the executive particular weight in treaty interpretation.

c. Congressional reassertions of power

Unaided by courts, senators have sought to reassert their role in the treaty-making process by participating in executive delegations to negotiate treaties. Senators were appointed to commissions to negotiate treaties as early as 1898,

75. 299 U.S. 304, 319 (1936).
76. Kolovrat v. Oregon, 366 U.S. 187, 194 (1961). The Court has also legitimated executive power to withdraw from treaties. In Goldwater v. Carter, 444 U.S. 996 (1979), the Court dismissed a Congressional challenge to a unilateral treaty termination by President Jimmy Carter, in effect legitimating the President’s conduct. The plaintiffs argued that the President’s actions “deprived them of their constitutional role with respect to a change in the supreme law of the land,” id. at 997-98 (Powell, J., concurring), but the Court dismissed the complaint as nonjusticiable. Id at 998. The Court’s decision was made without oral argument, just days before the scheduled termination. CONSTITUTIONAL CONFLICTS, supra note 12, at 238. Justice Brennan, the only justice to comment on the merits of the case, dissented from the Court’s decision to deny certiorari and indicated that he would have decided on the merits in the President’s favor. Goldwater, 444 U.S. at 1006 (Brennan, J., dissenting). Brennan noted “the President’s well-established authority to recognize, and withdraw recognition from, foreign governments” and found that the executive had authority to unilaterally terminate a treaty on that basis. Id. at 1007. Specifically, Brennan stated that

[a]brogation of the defense treaty with Taiwan was a necessary incident to Executive recognition of the Peking Government, because the defense treaty was predicated upon the now-abandoned view that the Taiwan Government was the only legitimate political authority in China. Our cases firmly establish that the Constitution commits to the President alone the power to recognize, and withdraw recognition from, foreign regimes. That mandate being clear, our judicial inquiry into the treaty rupture can go no further. Id. (citations omitted). Though Brennan’s opinion was not joined by any other justices, the effect of the court’s dismissal was to permit Carter’s withdrawal from the treaty. Thus, “[a]t the end of the twentieth century, it is apparently accepted that the President has authority under the Constitution to denounce or otherwise terminate a treaty.” HENKIN, supra note 9, at 214.
and the practice has increased in the last 60 years.\footnote{77. TREATIES AND OTHER INTERNATIONAL AGREEMENTS, supra note 73, at 109-10.} Four of eight U.S. representatives sent to the San Francisco Conference establishing the United Nations were members of Congress.\footnote{78. Id. at 110. This practice was controversial because, according to some, participation of members of Congress in executive delegations violates the incompatibility clause of the constitution. As the clause provides, \[no Senator or Representative shall, during the Time for which he was elected, be appointed to any civil Office under the Authority of the United States, which shall have been created, or the Emoluments whereof shall have been increased during such time; and no Person holding any Office under the United States, shall be a Member of either House during his Continuance in Office.\] U.S. CONST. art. I, § 6, cl. 2. Some members of Congress previously resigned their seats in order to serve on presidential delegations. TREATIES AND OTHER INTERNATIONAL AGREEMENTS, supra note 73, at 110. However, serving on executive delegations is more common at present so that the practice is no longer controversial. See Henkin, supra note 9, at 178.} In negotiating the North Atlantic Treaty, Secretary of State Dean Acheson relied heavily on the advice of Senators Thomas Connally and Arthur Vandenberg.\footnote{79. DEAN ACHESON, PRESENT AT THE CREATION: MY YEARS AT THE STATE DEPARTMENT 277 (1969).} Though the ability of members of Congress to influence particular provisions of a final agreement is unclear, Acheson wrote that, in speaking with his counterparts from other countries in the North Atlantic Treaty negotiations, “I pointed out the unwisdom of our trying to move any faster with a text than I could move in my discussions with Senators Connally and Vandenberg, since agreement would mean little unless it carried a senatorial opinion behind it.”\footnote{80. Id.}

In addition, in the last fifty years, the Senate has reasserted its power in international agreements through frequent use of conditional consent.\footnote{81. The Senate also used conditional consent to shape U.S. obligations in early treaties. See CURTIS A. BRADLEY & JACK L. GOLDSMITH, FOREIGN RELATIONS LAW 435-36 (2d ed. 2006) (discussing conditions on the Senate’s consent to the Jay Treaty of 1795 and the Treaty of Amity, Commerce, and Navigation signed by the U.S. and Tunisia in 1797).} Rather than ratifying treaties exactly as submitted by the President, the Senate now issues its consent subject to ex-post conditions, called reservations, understandings, or declarations (collectively RUDs).\footnote{82. Henkin, supra note 9, at 180.} RUDs may define treaty obligations for domestic purposes, reserve U.S. consent to certain obligations, or clarify the meaning of treaty terms to the Senate’s satisfaction.\footnote{83. Id. at 180-82.} Conditional consent has even been used to limit the President’s authority in treaty interpretation.\footnote{84. The so-called Biden Condition attached to the Treaty on the Elimination of Their Intermediate Range and Shorter Range Missiles made the Senate’s advice and consent subject to an understanding that “[t]he United States shall interpret the treaty in accordance with the common understanding of the treaty shared by the President and the Senate at the time the Senate gave its advice and consent to ratification.” 134 Cong. Rec. S6724 (daily ed. May 26, 1988).} Conditional consent has become common in treaty practice and
has generally facilitated U.S. consent to treaties. More importantly, this limit on executive power in treaties seems to have been successful: Courts generally give effect to RUDs, and even the executive branch now concedes that conditions on the Senate’s consent “are considered to be part of the United States’ position in ratifying the treaty, [and] they are generally binding on the President, both internationally and domestically, in his subsequent interpretation of the treaty.”

2. Alternative forms of international agreements

In addition to allowing treaty negotiation by the executive alone, U.S. practice evolved to permit the formation of international agreements without the approval of two-thirds of the Senate. Most importantly, the President now regularly concludes congressional executive agreements, which are negotiated by the President and approved by a simple majority of both houses.

a. Presidential assertions of power

Though early congressional executive agreements were made with prior congressional authorization, the President eventually began initiating negotiations for international agreements himself. As early as 1792, agreements for the development of an international postal service provided that “the Postmaster General may make arrangements with the postmasters in any foreign

85. RUDs have also affected international agreements outside the human rights context. For example, in the 1970s, the President sought Senate consent to a treaty with Spain on military bases that purported to mandate the appropriation of funds for the treaty over five years. CONSTITUTIONAL CONFLICTS, supra note 12, at 233. Both houses objected to this provision because it would undermine the authority of the Senate Committee on Foreign Relations and the House Committee on Foreign Affairs. Id. The Senate eventually ratified the treaty, but subject to a declaration that funds for the treaty “shall be made available for obligation and through the normal procedures of the Congress, including the process of prior authorization and annual appropriations.” Id. (quoting treaty language). When the military base treaty was replaced by an executive agreement in 1982, limitation on the appropriation of funds remained intact. Id.


87. BRADLEY & GOLDSMITH, supra note 81, at 448 (quoting Memorandum from Charles J. Cooper, Assistant Attorney General, Office of Legal Counsel, to Abraham D. Sofer, Legal Adviser, Department of State, “Relevance of Senate Ratification History to Treaty Interpretation,” 11 Op. Off. Legal Counsel 36 (1987)).

88. Executive practice has also developed the use of sole executive agreements, which are concluded by the President without Congressional approval. The President’s ability to make sole executive agreements is, however, subject to limitations. HENKIN, supra note 9, at 222. The President may conclude such agreements to govern diplomatic relations, settlement of claims, and military matters related to the Commander in Chief powers, but sole executive agreements may not commit the United States to war or to appropriate money. Id. at 229. Because sole executive agreements are not seen as largely interchangeable with treaties, they are not discussed in this article.
country for the reciprocal receipt and delivery of letters and packets, through the post-offices.”89 This authority was expanded in the McKinley Tariff Act of 1890, which authorized the President to negotiate subsequent reciprocal trade agreements to lower tariffs.90 The agreements negotiated by the President “looked so much like international treaties that they were frequently referred to as ‘treaties,’ even though they were never submitted to the Senate.”91 By authorizing the President to conclude these reciprocal trade agreements,

Congress also repudiated the original vision of the Treaty Clause. The clause was, after all, intended precisely to make it difficult, if not impossible, for the federal government to enter into international agreements . . . that would benefit one part of the country at the expense of another. By . . . authorizing the President to enter into trade agreements that would do exactly that, Congress turned its back on this original vision. The President could enact trade policy that he believed to be in the national interest, even if doing so might hurt manufacturers centered in particular states. And he could do so with the concurrence of a simple majority of both houses of Congress instead of a supermajority of the senate.92

Gradually, the executive began to exercise this power outside the trade context and to initiate negotiations independently, seeking congressional approval through a statute or other means only after concluding an agreement.93 These “ex post” congressional executive agreements took an even stronger hold after the Versailles Treaty failed to pass the Senate in the traditional treaty-making format.94 Though congressional executive agreements are now used in many areas of presidential authority, they are still most frequently used in trade. Most recently, the North American Free Trade Agreement was concluded by congressional executive agreement, rather than by treaty.

The substitution of congressional executive agreements for treaties constitutes a shift of power to the executive because congressional executive agreements do away with the supermajority requirement for treaty-approval in the Senate. Because treaties and congressional executive agreements are now

89. Act of Feb. 20, 1792, ch. 7 § 26, 1 Stat. 233, 239.
91. Hathaway, supra note 69, at 1295. Other statutes included similar provisions. For example, the Reciprocal Trade Agreements Act, ch. 474, 48 Stat. 943 (codified as amended at 19 U.S.C. § 1351 (2000)), permitted the President to negotiate bilateral agreements to reduce reciprocal tariffs up to fifty percent.
92. Hathaway, supra note 69, at 1298.
93. Id. at 1298-99.
94. Id. at 1299-1300. Some attribute the Senate’s failure to ratify the Versailles Treaty to pure partisanship. Henkin, supra note 9, at 178-79. Others assert that the President’s decision to negotiate the treaty unilaterally, without Senate advice, accounts for the Senate’s failings in this regard. Constitutional Conflicts, supra note 12, at 223-24.
seen as largely interchangeable, this constitutional requirement is no longer effective, even outside the trade context.

b. Judicial validation

Though the Supreme Court has not expressly authorized the formation of executive agreements, it has assumed their validity in several cases. Specifically, in *Field v. Clark*, the Court upheld a congressional delegation of tariff-setting authority to the President based partially on Congress’s history of permitting the President to negotiate executive agreements on trade. The use of congressional executive agreements is now so well-established that plaintiffs challenging NAFTA simply “concede[d] the existence of some kinds of valid and binding international agreements that do not constitute Article II treaties.” In that case, a federal district court upheld the use of an executive agreement to conclude NAFTA, noting “[i]t is impossible to definitively conclude that the Framers intended the regulation of foreign commerce to be subject to the rigors of the Treaty Clause procedure when commercial agreements with foreign nations are involved.” Thus, the use of powers outside the text of the Constitution have been approved in the context of foreign commerce.

c. Congressional reassertions of power

Though the President had been concluding congressional executive agreements for almost 200 years, Congress sought to reassert its power over non-treaty international agreements in 1972 by requiring notice. The Case Act of

---

96. BRADLEY & GOLDSMITH, supra note 81, at 475.
98. Made in the USA Found. v. United States, 56 F. Supp. 2d 1226, 1292 (N.D. Ala. 1999), vacated by 242 F.3d 1300 (11th Cir. 2001).
99. Id. at 1319-20.
100. Though courts have invalidated executive agreements rather than challenge their basic credibility in other contexts, these cases establish the permissible boundaries of these agreements. In *Seery v. United States*, 127 F. Supp. 601 (Ct. Cl. 1955), the Court of Claims struck down an executive agreement purporting to settle all obligations incurred by U.S. armed forces in Austria during World War II, finding that it constituted a deprivation of property without due process of law. As the court stated, “there can be no doubt that an executive agreement, not being a transaction which is even mentioned in the constitution, cannot impair Constitutional rights.” Id. at 606. The Supreme Court also struck down an executive agreement in *Reid v. Covert*, 354 U.S. 1, 16-20 (1957), holding that American servicemen and their dependents tried abroad could not be deprived of their Fifth and Sixth Amendment rights at trial by executive agreement. Finally, the Fourth Circuit has also struck down an executive agreement because it contravened an existing commercial statute. United States v. Guy W. Capps, Inc., 204 F.2d 655, 658-60 (4th Cir. 1953), aff’d on other grounds, 348 U.S. 286 (1955). Though these decisions limited executive power over international agreements in individual cases, they actually served to legitimize executive power in the long run by affirming that executive agreements are generally permissible, so long as they do not cross certain boundaries.
2009] FRANKFURTER’S GLOSS THEORY 123

1972 requires the Secretary of State to “transmit to the Congress the text of any international agreement . . . other than a treaty” within sixty days of the agreement’s entry into force.\textsuperscript{101} Agreements whose “immediate public disclosure. . . would, in the opinion of the President, be prejudicial to the national security of the United States,” must be transmitted “under an appropriate injunction of secrecy” to the Senate Committee on Foreign Relations and the House Committee on International Relations rather than the full Congress.\textsuperscript{102}

Executive compliance with these reporting requirements, however, is questionable. A 1976 study by the Government Accountability Office revealed that Presidents had entered into a number of arrangements with foreign powers without submitting them to Congress or even to the State Department’s Office of Treaty Affairs.\textsuperscript{103} Further, according to a 1977 study, almost forty percent of executive agreements concluded in the previous year had been submitted to Congress after the sixty-day window supplied in the Case Act.\textsuperscript{104} Despite congressional attempts to remedy these problems with the Case Act,\textsuperscript{105} executive compliance remains questionable.\textsuperscript{106}

3. Critical analysis

Therefore, in international agreements, Frankfurter’s gloss on executive power has largely shifted authority to the President at the expense of Congress.

\textsuperscript{102} Id.
\textsuperscript{103} CONSTITUTIONAL CONFLICTS, supra note 12, at 246-47.
\textsuperscript{104} Id. at 247.
\textsuperscript{105} In 1977, Congress amended the Case Act to require all U.S. government agencies and departments to transmit all agreements concluded on behalf of the United States to the State Department. Id. Congress also broadened the definition of “executive agreement” in 1978 to include oral agreements that “shall be reduced to writing.” Id. Finally, in 1987 Congress passed a bill providing that it would not appropriate funds to implement any agreement that had not been properly transmitted under the Case Act within the 60 day window until that agreement had been transmitted to Congress. Foreign Relations Authorization Act, Fiscal Years 1988 and 1989, Pub. L. No. 100-204, § 139, 101 Stat. 1331, 1347 (1987).
\textsuperscript{106} Presidents continue to sidestep the Case Act’s reporting requirements by issuing “parallel policy statements” and “agreed frameworks” with foreign leaders rather than concluding formal agreements that would require reporting. For example, in 1977, ten days before the scheduled expiration of SALT I, U.S. and Soviet officials issued separate statements of intent to adhere to the treaty. CONSTITUTIONAL CONFLICTS, supra note 12, at 247. Similarly, in 1994, the United States and North Korea concluded an “agreed framework” for modernization of Korea’s reactors and to prevent North Korea from developing nuclear weapons. Id. Though President Clinton’s addresses to North Korea contained language of commitment, the administration officially considered the framework a nonbinding political agreement and therefore did not report it under the Case Act. Id. at 247-48. In addition, because it still sees a distinction between treaties and executive agreements, the Senate has declined to consent to the Vienna Convention on the Law of Treaties, which treats both as “international agreements.” TREATIES AND OTHER INTERNATIONAL AGREEMENTS, supra note 73, at 20-21.
In the formation of international agreements, the legislature initially ceded power to the executive due to the Senate’s ineffectiveness and lack of discretion in providing advice and consent in the treaty-making process. Congress regained some power over the negotiation of international agreements through the participation of individual legislators in executive delegations and by attaching ex-post conditions on the Senate’s consent. Executive officials have acknowledged the role of individual senators in the negotiation of international agreements, and both courts and the executive consider conditions on the Senate’s consent to treaties as binding.

With respect to the form of international agreements, politicized decision-making by the Senate caused the President to abandon the use of treaties for some matters, and instead to use executive agreements. Because Congress did not seek to reassert its power until executive practices were well-established to permit the conclusion of non-treaty international agreements, the Case Act of 1972 has not proven ineffective at controlling presidential conduct.

B. War

Unlike the treaty power, war powers are included in both Article I and Article II of the Constitution. Under Article I, congressional war powers include the power

To declare War, grant Letters of Marque and reprisal, and make Rules concerning Captures on Land and Water; To raise and support Armies . . . ; To provide and Maintain a Navy; To provide for calling forth the Militia to execute Laws of the Union, suppress Insurrections and repel Invasions; [and] To provide for organizing, arming, and disciplining, the Militia, and for governing such Part of them as may be employed in the Service of the United States.107

Article II provides “[t]he President shall be Commander in Chief of the Army and Navy of the United States.”108 Though the bulk of war powers assigned in the Constitution are granted to Congress, the President has asserted broad powers in war under Frankfurter’s gloss theory, as in international agreements.

1. Presidential assertions of power

Presidential assertions of defensive war powers began with a congressional statute. In 1795, Congress delegated to the President the power to call forth the militia and to issue orders that he felt proper “whenever the United States shall be invaded, or be in imminent danger of invasion from any foreign nation or Indian tribe.”109 Because the President was expected to unilaterally respond in

kind to such attacks, the President’s authority under this statute has expanded as weapons have grown more powerful. Therefore, “during the decades of the Cold war, everyone . . . assumed that the President would have power also to retaliate against a nuclear attack.”

The President has also frequently used U.S. armed forces for purposes short of war without congressional approval. Since such power has been exercised repeatedly by Presidents of both major political parties, and at times with congressional and public support, even members of Congress now commonly accept that the President may deploy the U.S. military “for foreign policy purposes, even to court or engage in hostilities short of war” without a declaration of war. Until the 1950s, however, Presidents generally deployed the military with congressional approval, even if not in the form of a declaration of war. For example, Congress did not declare war in the “Quasi War” with France from 1798 to 1800, but it passed statutes increasing the size of the military and reinforcing port defense, therefore approving the President’s use of force indirectly. Since the United Nations Charter prohibited the use of force in international law, congressional declarations of war have disappeared from U.S. practice. Instead, when Congress authorizes force, it does so through an “authorization of military force” or by providing funds. After World War II, Presidents took another step, asserting power to conduct extensive armed conflicts without congressional approval. Thus, Congress did not formally declare war in the Vietnam War, the Korean War, the Gulf War, or the Iraq War, thereby permitting the assertion of even broader unilateral executive authority in war.

2. Judicial validation

As in international agreements, courts have approved the executive’s assertions of power over war. The Supreme Court recognized the President’s power to repel invasions very early in U.S. history. As the Court declared in The Prize Cases, “[i]f a war be made by invasion of a foreign nation, the President is not only authorized but bound to resist force by force.” The Supreme Court upheld Congress’s delegation of defensive use of force to the President, and even expanded that power. After upholding the 1795 Act in Martin v. Mott, the Court held that “the authority to decide whether the exigency [warranting unilateral

110. HENKIN, supra note 9, at 48.
111. Id.
112. Id. at 99.
114. Id. at 15-16.
115. Id.
executive action] has arisen belongs exclusively to the President and . . . his decision is conclusive upon all other persons."

The Supreme Court has also upheld offensive uses of force by the executive without a prior congressional declaration of war. In *Bas v. Tingy*, the Court distinguished perfect from imperfect wars, finding that a general declaration of war is not required to legitimate more limited, imperfect wars.\footnote{118} In that case, Congress had passed a series of increasingly hostile measures toward France that finally culminated in all-out war. The court explained,

> As the temper of the people rose . . . the language and the measures of the government became more and more energetic and indignant; though hitherto the popular feeling may not have been ripe for a solemn declaration of war; and an active and powerful opposition in our public councils, has postponed, if not prevented that decisive event, which many though would have best suited the interest, as well as the honour of the United States.\footnote{119}

Thus, even though Congress had not declared a general war against France, the Court found that a state of war had in fact existed due to the successive measures passed by Congress indicating hostility toward France.

The Second Circuit espoused a similar view in the context of a much more extensive conflict. In *Orlando v. Laird*, the Second Circuit held that the Vietnam War was legitimate even though Congress had not formally declared war.\footnote{120} Further, the court found that if the Constitution did require a declaration of war, the Gulf of Tonkin Resolution would suffice.\footnote{121} Though the court reached the merits on the validity of the war itself, it held that the appropriate means of “declaring war” under the Constitution was a nonjusticiable political question.\footnote{122} As the court stated, “[t]he form which congressional authorization should take is one of policy, committed to the discretion of the Congress and outside the power and competency of the judiciary, because there are no intelligible and objectively manageable standards by which to judge such actions.”\footnote{123} In so holding, however, the court effectively found that the constitutional requirement of a declaration of war may be fulfilled by informal and indirect congressional authorization, even in large-scale conflicts like the Vietnam War.

\footnotesize{
117. 25 U.S. (12 Wheat.) 19, 30 (1827).
118. 4 U.S. (4 Dall.) 37 (1800).
119. Id. at 45.
120. 443 F.2d 1039, 1042-43 (2d Cir. 1971).
121. Id.
122. Id. at 1043.
123. Id.
}
3. Congressional reassertions of power

In the wake of the Vietnam War, Congress sought to reassert its control over war by passing the War Powers Resolution in 1973. The first section of the resolution purports to limit the President’s use of force as Commander in Chief to specific situations, but lacks binding language. As the resolution provides,

The constitutional powers of the President as Commander-in-Chief to introduce United States Armed Forces into hostilities, or into situations where imminent involvement in hostilities is clearly indicated by the circumstances, are exercised only pursuant to (1) a declaration of war, (2) specific statutory authorization, or (3) a national emergency created by attack upon the United States, its territories or possessions, or its armed forces.

The statute does use binding language with respect to reporting. Under the second section, the President must report any use of force without a declaration of war to congressional leaders within 48 hours of the introduction of U.S. forces into hostilities. The resolution also requires the President to report to Congress periodically and to provide any other information requested by Congress outside these reports. Perhaps most daringly, the War Powers Resolution requires the President to terminate use of the military if Congress does not approve the deployment within sixty days of receiving notice under the act.

However, the War Powers Resolution proved unsuccessful in binding the executive’s use of force. Presidents have deployed U.S. forces abroad at least fifteen times since the resolution was enacted, but congressional approval was attained only on three occasions. In the Gulf War, though President George H. W. Bush eventually received congressional approval to use force, he deployed U.S. forces to Saudi Arabia without even acknowledging any requirements under the War Powers Resolution. In fact, the War Powers Resolution has proven so ineffective that a bipartisan study group recently proposed new legislation to replace it.

---

126. Id. § 1541(c).
127. Id. § 1543(a).
128. Id. § 1543(b).
129. Id. § 1544(b).
130. BRADLEY & GOLDSMITH, supra note 81, at 253.
131. HENKIN, supra note 9, at 109-10.
132. National War Powers Commission, http://millercenter.org/policy/commissions/warpowers (last visited Apr. 1, 2009). The Commission includes former officials from both the executive and legislative branches. Id. Their proposal, called the War Powers Consultation Act of 2009, would require the President to consult with Congress before deploying military forces in a “significant armed conflict.” Id. The Commission defines the term “significant armed conflict” and would
Congress has been more successful in restricting presidential war powers on an individual basis. For example, Congress required the executive to withdraw troops from Somalia by prohibiting the Department of Defense from spending its funds there after March 31, 1994. \(^{133}\) Similarly, Congress provided conditions under which U.S. forces had to lift enforcement of the U.N. arms embargo against Bosnia. \(^{134}\) Similar to RUDs applied to international agreements negotiated by the President, these ex-post limitations on executive power have proven more effective than the broad ex-ante prohibitions in the War Powers Resolution.

4. Critical analysis

The operation of Frankfurter’s gloss theory on executive power in international agreements and war reveals three trends. First, in both areas, reporting requirements proved ineffective under Frankfurter’s gloss theory. The President has ignored or evaded such requirements included in both the Case Act and the War Powers Resolution, perhaps because neither statute provides a sanction for failure to comply. Second, the operation of Frankfurter’s gloss theory in these two areas demonstrates that courts generally favor the executive in foreign relations decisions. Though courts sometimes find in favor of Congress in individual cases, the long-run consequence of judicial decisions has been to carve out a sphere of legitimate executive authority over international agreements and war at Congress’s expense. Finally, Frankfurter’s gloss theory in these two areas has revealed the effectiveness of individual, ex-post regulation of presidential conduct. Just as RUDs provided as conditional consent to individual agreements have proven effective in shaping U.S. obligations in international agreements, congressional statutes limiting the executive’s power in individual uses of force have proven effective in actually limiting that authority. Despite the effectiveness of these individual, ex-post limitations, however, Frankfurter’s gloss theory has generally served to increase the President’s power over both war and international agreements.

C. Surveillance

In surveillance, as in international agreements and war, Frankfurter’s gloss theory on executive power has generally shifted authority to the President at the expense of Congress. Unlike international agreements and war, however, the


\(^{134}\) Pub. L. No. 103-160, Div. A Title XV § 1511, 107 Stat. 1839 (1993). In 1996, members of Congress proposed a broader limit on executive power—a prohibition on the use of Department of Defense funds for armed forces under U.N. control. For discussion and the reaction of the President’s lawyers, see Bradley & Goldsmith, supra note 81, at 251.
Constitution does not directly address surveillance, therefore making the history of separation of powers more decisive in the perceived legitimacy of congressional regulation.

1. Presidential assertions of power

Though the Constitution includes several provisions relevant to surveillance, that authority is not directly mentioned or assigned to any branch of government. For much of American history, therefore, this absence of constitutional authorization for surveillance power was interpreted in favor of the executive under the assumption that surveillance is an executive authority that is vested with the President under Article II of the Constitution. Foreign intelligence-gathering was traditionally seen as “an extension of [the President’s] role as ‘sole organ’ and his traditional function as ‘the eyes and ears’ of the United States.” Therefore, “Congress consistently deferred to the President when he withheld secret official records, when he employed secret agents, when he ransomed hostages, and even when he engaged in covert operations.”

Specifically, for the first 150 years of American history, Congress funded the President’s foreign intelligence activities and withdrew requests for information when the executive indicated that disclosure would interfere with national security. Because “the nature of intelligence activities rarely touched the private lives of the citizenry” during this time, presidential power faced neither popular nor political resistance.

2. Judicial validation

As in international agreements and war, courts generally approved broad executive powers in foreign surveillance, and even recognized executive powers in foreign surveillance while limiting law enforcement in the domestic context. In Katz v. United States, the Supreme Court applied the Fourth Amendment’s warrant requirement to domestic wiretapping, but noted that its analysis did not

135. In particular, the First and Fourth Amendments have served as limits on surveillance authority. Under the First Amendment, “Congress shall make no law . . . abridging the freedom of speech, or of the press; or the right of the people to peaceably to assemble.” U.S. CONST. amend I. As the Fourth Amendment provides, “[t]he right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause.” U.S. CONST. amend IV.


137. HENKIN, supra note 9, at 111.


139. Id.

140. Id. at 17.
apply to national security wiretaps. 141 Similarly, in the Keith case, the Court struck down domestic wiretapping for national security purposes without prior judicial approval, but again declined to apply its analysis to foreign surveillance. 142 In a footnote, however, the Court stated that “warrantless surveillance, though impermissible in domestic security cases, may be constitutional where foreign powers are involved.” 143

After the Supreme Court declined to rule definitively on foreign surveillance in these two cases, lower courts granted the executive greater power. The Fourth Circuit recognized the executive’s inherent foreign surveillance power in United States v. Truong Dinh Hung, stating that “the executive possesses unparalleled expertise to make the decision whether to conduct foreign intelligence surveillance.” 144 Further, the Fourth Circuit found that “the executive branch not only has superior expertise in the area of foreign intelligence, it is also constitutionally designated as the pre-eminent authority in foreign affairs.” 145 The court drew a parallel to Keith, noting that “[j]ust as the separation of powers in Keith forced the executive to recognize a judicial role when the President conducts domestic security surveillance, so the separation of powers requires us to acknowledge the principal responsibility of the President for foreign affairs and concomitantly for foreign intelligence surveillance.” 146

Even the United States Foreign Intelligence Surveillance Court of Review, a court created by Congress to limit the executive’s surveillance power, has deferred to the executive. In its sole opinion, the Court of Review affirmed the Fourth Circuit’s inherent executive powers holding, noting that “[t]he Truong court, as did all the other courts to have decided the issue, held that the President did have inherent authority to conduct warrantless searches to obtain foreign intelligence information . . . We take for granted that the President does have that authority.” 147 Therefore, the court reasoned, Congress could not pass statutes that would “encroach on the President’s constitutional power.” 148 Thus, in the

141. 389 U.S. 347, 358 n.23 (1967). In his concurring opinion, Justice White went further than the majority, stating “[w]e should not require the warrant procedure and the magistrate’s judgment if the President of the United States or his chief legal officer, the Attorney General, has considered the requirements of national security and authorized electronic surveillance as reasonable.” Id. at 364 (White, J., concurring). Justices Douglas and Brennan, however, disagreed with this approach, calling it a “wholly unwarranted green light for the Executive Branch to resort to electronic eavesdropping without a warrant in cases which the Executive Branch itself labels ‘national security’ matters.” Id. at 359 (Douglas, J., concurring).

142. United States v. U.S. Dist. Court, 407 U.S. 297, 321-22 (1972). As the Supreme Court stated, “[w]e . . . express no opinion as to, the issues which may be involved with respect to activities of foreign powers or their agents.” Id.

143. Id. at 322 n.20.

144. 629 F.2d 908, 913 (4th Cir. 1980) (citations omitted).

145. Id. at 914.

146. Id. (citing United States v. U.S. Dist. Court, 407 U.S. 297, 321-22 (1972)).

147. In re Sealed Case No. 02-001, 310 F.3d 717, 742 (FISA Ct. Rev. 2002) (citing United States v. Truong Dinh Hung, 629 F.2d 908 (4th Cir. 1980)).

148. Id.
absence of Supreme Court precedent, lower federal courts have granted the executive broad discretion in foreign intelligence surveillance, declining to apply the Fourth Amendment safeguards that restrict domestic surveillance.

3. Congressional reassertions of power

As technology advanced and greater surveillance capabilities threatened the privacy of ordinary citizens, Congress took a more active role in the executive’s surveillance functions. As evidence of intelligence abuses came to light, Congress expanded its role in overseeing and investigating the intelligence community by establishing the Commission on CIA Activities Within the United State (the Rockefeller Commission) and the Select Committee to Study Governmental Operations With Respect to Intelligence Activities (the Church Committee) in 1975.149 After finding abuses and inefficiencies in the executive’s foreign surveillance methods, the Church Committee recommended the passage of legislation to improve surveillance methods and to limit the exercise of executive authority.150

Three years later such regulation began with the enactment of the Foreign Intelligence Surveillance Act of 1978 (FISA).151 FISA created a procedure to permit the executive branch to obtain special warrants for wiretapping and physical searches of foreign entities. To hear applications for these warrants, FISA established the Foreign Intelligence Surveillance Court (FISA Court) and the Foreign Intelligence Surveillance Court of Review.152 If the executive shows that “the target of the electronic surveillance is a foreign power or an agent of a foreign power” and that “each of the facilities or places at which the electronic surveillance is directed is being used, or is about to be used, by a foreign power or an agent of a foreign power” the statute requires the FISA court to issue a warrant.153 In 1988, Congress added a provision to FISA, indicating that this procedure was “the exclusive means” for conducting electronic surveillance.154

---

149. Banks & Bowman, supra note 139, at 32-35. The Rockefeller Commission established that the purpose of surveillance was the relevant criteria for determining the validity of executive surveillance. While warrantless surveillance for the purpose of criminal prosecution is prohibited, warrantless surveillance for national security purposes is permissible. Id. at 33. This distinction was struck down by the FISA Court of Appeals in In re: Sealed Case No. 02-001, 310 F.3d 717 (FISA Ct. Rev. 2002).

150. Banks & Bowman, supra note 139, at 34.


152. 50 U.S.C. § 1803.

153. Id. § 1805 (emphasis added).

According to the executive branch, however, the President retained his inherent surveillance authority even after FISA was enacted.\textsuperscript{155} After the attacks of September 11, 2001, the President secretly authorized the National Security Agency (NSA) to eavesdrop on phone calls placed to foreign locations from the United States without obtaining FISA warrants.\textsuperscript{156} Though congressional leaders and the Chief Judge of the FISA court were notified of the program, it was not publicly reported until December 15, 2005, and was never approved under FISA procedures.\textsuperscript{157} Administration officials defended the program as part of the President’s inherent powers and as authorized under FISA. According to the Department of Justice’s memorandum to Congress in defense of the NSA program,

> Article II of the Constitution vests in the President all executive power of the United States, including the power as Commander in Chief of the Armed Forces . . . and authority over the conduct of the Nation’s foreign affairs . . . To carry out these responsibilities, the President must have authority to gather information necessary for the execution of this office . . . Thus, it has long been recognized that the President has the authority to use secretive means to collect intelligence necessary for the conduct of foreign affairs and military campaigns . . . In reliance on these principles, a consistent understanding has developed that the President has inherent constitutional authority to conduct warrantless searches and surveillance within the United States for foreign intelligence purposes. Wiretaps for such purposes thus have been authorized by Presidents at least since the administration of Franklin Roosevelt in 1940.\textsuperscript{158}

Thus, in the executive’s view at least, FISA does not constitute the sole means of authorized surveillance by the federal government.

Though the NSA program provoked controversy in the media and the legal community,\textsuperscript{159} Congress subsequently authorized the executive to conduct a

---


\textsuperscript{157} \textit{Id.}


\textsuperscript{159} A number of former government lawyers turned legal academics joined to write letters to members of Congress expressing their concerns about the NSA program and directly rebutting the government’s defense of the program. One letter is reprinted in \textit{On NSA Spying: A Letter to Congress}, \textit{53 N.Y. Review of Books} 2 (Feb. 9, 2006), \textit{available at} http://www.nybooks.com/articles/18650.
similar program without FISA warrants for a limited period of time. The Protect America Act, which expired on February 15, 2008, permitted the Attorney General and the DNI to use reasonable procedures to “authorize the acquisition of foreign intelligence information concerning persons reasonably believed to be outside the United States” if “a significant purpose of the acquisition is to obtain foreign intelligence information.” Eventually, the executive agreed to keep its surveillance within FISA limitations, but maintained that it had inherent authority to conduct surveillance outside the FISA procedures if necessary to protect national security.

After protracted debate, Congress responded by passing the FISA Amendments Act of 2008. Rather than reasserting Congress’s role in surveillance, this act actually reneged on some previous attempts to reign in executive surveillance. The 2008 amendments permit the DNI and the Attorney General to order surveillance for up to one year without the approval if a FISA court grants immunity to phone companies that previously permitted the NSA to eavesdrop on the conversations of their customers. Therefore, though the Act does indicate Congress is still active in surveillance, its substantive provisions do not actually serve to strongly limit the power of the executive in that area.

4. Critical analysis

As a whole, Congress has ceded surveillance authority to the President. As in international agreements, the shift of power to the President in surveillance is primarily due to Congress’s inaction and ineffectiveness. By failing to legislate on the permitted scope of foreign intelligence surveillance by the executive, Congress permitted the President to establish a long-standing practice of surveillance apparently based on his inherent authority. When Congress did legislate, its limitations on executive power proved ineffective because long-standing unilateral surveillance by the executive led to assertions of inherent authority to do so.

Further, the particular limits Congress chose to enact were ill-advised based on previous experience. In light of judicial decisions favoring the executive in international agreements and war, establishing a court to check executive surveillance authority was unlikely to prove effective. In addition, the fact that the FISA court must approve the executive’s surveillance decisions if certain

162. 50 U.S.C. § 1881(q).
criteria are satisfied does not allow judges sufficient latitude to check executive surveillance. A more effective means of regulating foreign intelligence surveillance by the executive may have been through legislative oversight, rather than judicial authorization. If judicial oversight was used, however, Congress should have granted FISA judges greater authority to approve or reject executive actions in surveillance.

On the other hand, congressional regulation in surveillance offers a significant advantage over congressional legislation in war. Whereas the War Powers Resolution seeks to limit executive authority in non-binding terms, Congress makes more affirmative statements of its authority under FISA to establish the exclusive means of surveillance. Though the executive has engaged in surveillance outside FISA, it has also heeded FISA’s limits more readily than the limits of the War Powers Resolution. Whereas the executive has generally refused to even acknowledge the applicability of the War Powers Resolution, the executive usually recognizes FISA’s validity and simply asserts that the statute in fact permits the executive action in question.

III. HISTORICAL FRAMEWORK FOR SEPARATION OF POWERS

As demonstrated above, the history of separation of powers in international agreements, war, and surveillance establishes a framework of factors weighing for and against a shift of power to the executive under Frankfurter’s gloss theory. Though the impact of individual factors has differed in each area of law, a consistent trend emerges when all three are examined. In international agreements, war, and surveillance, Presidents have successfully increased their powers based on congressional inaction, judicial approval, and ineffective congressional action. On the other hand, Congress has been more successful in maintaining or reasserting its role in foreign affairs when members of Congress can issue ex-post comments on individual actions, when legislators directly participate in the executive’s decisions, when statutes regulating the executive use mandatory language, and when the executive’s conduct is subject to media scrutiny. In addition, constitutional language on the power at stake is relevant as an initial matter to the perceived legitimacy of any congressional action.

A. Threshold consideration: constitutional language

As a threshold matter, constitutional language is relevant to the operation of Frankfurter’s gloss theory of executive power. Where the Constitution specifically envisions legislative participation, as in international agreements and war, expansion of executive power is generally more difficult. In these two areas, the Constitution provides a basis for the legitimate regulation of

165. See supra notes 68-106 and accompanying text.
166. See supra notes 107-134 and accompanying text.
executive action by the legislature, even when the initial understanding of these provisions has been abandoned in practice. On the other hand, where the Constitution provides no guidance and powers are presumed to be vested in the executive, as with surveillance,\textsuperscript{167} any congressional regulation whatsoever may be challenged as unconstitutional. Though some assert that specific congressional limits on executive power in international agreements and war are unconstitutional, blanket assertions of unconstitutionality have not been made in these areas.

Constitutional language, however, is not dispositive in determining the relationship between the executive and legislature in foreign relations. Though the Constitution requires the advice and consent of a two-thirds majority in the Senate for the formation of treaties, the executive has taken sole control over treaty negotiations and concludes other forms of international agreements without such Senate approval.\textsuperscript{168} Similarly, in war powers, though the Constitution requires a declaration of war, the executive has used armed forces frequently without such a declaration, citing the Commander in Chief powers.\textsuperscript{169} Therefore, even where the Constitution specifically envisions legislative participation, Congress must take positive steps to preserve its role.

B. Factors for shifting power to the executive

Regardless of constitutional language, Congress has ceded foreign relations power by failing to react following an assertion of power by the President. Specifically, in surveillance, decades of executive action with legislative acquiescence led to broad claims of inherent executive authority, which made the passage and enforcement of FISA more difficult.\textsuperscript{170} In war, because Congress has repeatedly permitted the deployment of U.S. forces without a congressional declaration of war, the Commander in Chief power is now consistently interpreted more broadly to permit such uses of force.\textsuperscript{171} Therefore, to preserve its power, Congress should act quickly after the executive begins asserting a new foreign relations power in order to curb claims of inherent authority or broad interpretations of constitutional grants of executive power.

Further, absent congressional action, courts have deferred to the executive in cases implicating foreign relations issues.\textsuperscript{172} Courts have found for the President on the merits or based on abstention doctrines in cases relating to the legality of deploying U.S. forces without a congressional declaration of war,\textsuperscript{173} the existence of treaty-like relationships based on agreements reached without the

\textsuperscript{167} See supra notes 135-164 and accompanying text.
\textsuperscript{168} See supra notes 88-106 and accompanying text.
\textsuperscript{169} See supra notes 107-134 and accompanying text.
\textsuperscript{170} See supra notes 151-164 and accompanying text.
\textsuperscript{171} See supra notes 115-134 and accompanying text.
\textsuperscript{172} See supra notes 141-148 and accompanying text.
\textsuperscript{173} See supra notes 116-123 and accompanying text.
Senate’s advice and consent,\(^{174}\) and the existence of inherent executive authority over surveillance.\(^ {175}\) As the decisions of the FISA courts demonstrate, even courts created for the specific purpose of limiting executive discretion tend to defer to the executive.\(^ {176}\) Therefore, when seeking to preserve its foreign relations power, Congress should not pass legislation that depends on courts to limit the foreign relations power of the executive.

Finally, ineffective congressional action in foreign relations has also allowed the executive branch to exercise broader authority. In international agreements, when ex ante “advice” from the Senate proved ineffective in the Washington administration, Congress failed to correct the process and provide a workable solution that permitted effective congressional participation in treaty-making.\(^ {177}\) A similar problem arose in surveillance when Congress charged courts with limiting executive abuses.\(^ {178}\) Therefore, ineffective congressional action, like congressional inaction and reliance on judicial oversight, has led to expanded executive power under Frankfurter’s gloss theory.

C. Factors against shifting power to the executive

The operation of Frankfurter’s gloss theory in international agreements, war, and surveillance, however, also demonstrates several factors that have helped Congress reassert or maintain its role in foreign relations. First, permitting Congress to comment on individual executive actions has effectively curtailed executive dominance of international agreements and war. Though international agreements are now negotiated by executive officials alone, the Senate maintains authority by attaching specific conditions to its consent to individual treaties that shape international obligations of the United States.\(^ {179}\) Similarly, though the War Powers Resolution has not limited the President’s unilateral use of force across the board, Congress has been able to control particular conflicts by setting conditions on the use of force or setting limits on military funding with respect to a particular conflict.\(^ {180}\)

In addition, direct participation by members of Congress in executive decision-making bodies has proven effective in maintaining or re-establishing the legislature’s role in foreign relations.\(^ {181}\) When members of Congress began participating in the executive’s treaty negotiation delegations, executive officials acknowledged the importance of their role, and the legislative branch was able to

\(^{174}\) See supra notes 96-100 and accompanying text.
\(^{175}\) See supra notes 141-148 and accompanying text.
\(^{176}\) See supra notes 147-148 and accompanying text.
\(^{177}\) See supra notes 89-95 and accompanying text.
\(^{178}\) See supra notes 149-164 and accompanying text.
\(^{179}\) See supra notes 81-87 and accompanying text.
\(^{180}\) See supra notes 133-134 and accompanying text.
\(^{181}\) See supra notes 78-80 and accompanying text.
re-establish its role in the formation of international agreements.\textsuperscript{182} As demonstrated by the failures of the War Powers Resolution\textsuperscript{183} and the Case Act,\textsuperscript{184} however, reporting requirements alone are insufficient to preserve separation of powers in foreign relations. Presidents have generally ignored such requirements, arguing that they violate inherent executive powers.

Statutes with mandatory language are also more effective in controlling the executive. For example, FISA’s explicit statement that the statute creates the exclusive constitutional means of conducting surveillance\textsuperscript{185} has preserved Congress’s role in foreign affairs more effectively than the precatory language of the War Powers Resolution.\textsuperscript{186} With respect to surveillance, though the executive has disputed this term and at times has violated it, Presidents have consistently acknowledged FISA.\textsuperscript{187} By contrast, Presidents have largely ignored the first section of the War Powers Resolution, presumably because it uses only precatory language.\textsuperscript{188} In fact, the War Powers Resolution has proven so ineffective that policymakers now advise completely replacing it with another law.\textsuperscript{189} On the other hand, because the initial framework of FISA was effective, Congress has only sought to amend the statute.\textsuperscript{190}

Finally, media attention has proven helpful in reestablishing or maintaining separation of powers in foreign relations. With respect to surveillance, the NSA spying story and subsequent debates were extensively covered in the media.\textsuperscript{191} In this coverage, the media exposed the excesses of executive power, and stirred significant public criticism of the NSA program.\textsuperscript{192} Since the executive’s decision to go beyond FISA were so strongly criticized in the media, future Presidents may be reluctant to make similar decisions for fear of such criticism.

\section*{IV. APPLYING THE FRAMEWORK TO FOREIGN INVESTMENT}

Applying the framework above to foreign investment reveals that FINSA and its precursors have been relatively successful in preserving separation of powers in foreign investment. However, this analysis also indicates several weaknesses in the current legislation that should be addressed in future amendments to assure Congress’s role in the area. As in other areas of foreign relations law, constitutional provisions are relevant as a threshold matter.

\begin{footnotes}
\footnotetext{182}{Id.}
\footnotetext{183}{See supra notes 124-133 and accompanying text.}
\footnotetext{184}{See supra notes 101-106 and accompanying text.}
\footnotetext{185}{18 U.S.C. § 2511.}
\footnotetext{186}{See supra notes 124-132 and accompanying text.}
\footnotetext{187}{See supra notes 155-164 and accompanying text.}
\footnotetext{188}{See supra notes 124-132 and accompanying text.}
\footnotetext{189}{See supra note 132 and accompanying text.}
\footnotetext{190}{See supra notes 163-164 and accompanying text.}
\footnotetext{191}{See supra notes 155-162 and accompanying text.}
\footnotetext{192}{Id.}
\end{footnotes}
A. Threshold consideration: constitutional language

As in surveillance, the Constitution does not specifically assign the regulation of foreign investment to the President or to Congress. Because Article I grants Congress the power “[t]o regulate Commerce with foreign Nations,” constitutional authority over foreign investment lies most reasonably in the hands of Congress. If regulation of foreign investment is regarded solely as a matter of foreign commerce, then congressional action would not face the type of blanket allegations of illegitimacy asserted with respect to FISA. In fact, executive action without congressional approval could be challenged as unconstitutional.

However, as the operation of Frankfurter’s gloss theory in international agreements and war has proven, constitutional language is not dispositive. Other constitutional powers could be interpreted to allow a broader, perhaps unilateral role for the executive in investment. Specifically, the President’s power to repel invasions, as acknowledged by Congress and affirmed in *The Prize Cases*, could serve as a basis for unilateral executive action in foreign investment. If foreign investment is regarded as an attack, the President’s power to repel invasions could be construed to permit executive control of foreign investment beyond the confines of FINSA. Because current laws already deem financial support to terrorist organizations a threat to the United States and make such support punishable as a crime, characterizing foreign investment as an invasion of the United States is conceivable. In framing foreign investment legislation, Congress should be careful to avoid the impression of declaring investment to be a form of attack, therefore avoiding the inadvertent creation or acknowledgement of unilateral executive authority to block transactions involving U.S. national security assets.

B. Review of current legislation

Though Congress has not specifically responded to these constitutional concerns, it has been relatively successful in preventing an expansion of executive power in foreign investment by remaining active in the area, preventing judicial legitimation of executive authority, and responding effectively to new concerns.

1. Congressional inaction

Unlike its legislation on international agreements, war, and surveillance, Congress legislated in foreign investment soon after the President began

194. Providing material support to terrorists is a crime under 18 U.S.C.A. § 2339A (West 2008). “Material support” includes “currency or monetary instruments or financial securities, [or] financial services.” *Id.*
exercising authority in the area. IEEPA, Congress’s first foreign investment statute, was enacted just two years after the Executive Order establishing CFIUS. By contrast, FISA was enacted after the President had been conducting surveillance without statutory authorization for 150 years. Further, Congress amended foreign investment legislation when necessary to ensure its effectiveness in responding to the federal government’s needs. When IEEPA proved ineffective because it applied only in cases of declared emergencies, Congress acted again four years later by passing the Exon-Florio provisions, establishing a concrete process for reviewing foreign investment in the United States for use outside declared emergencies. By acting quickly to regulate the area of foreign investment and by amending ineffective legislation, Congress may have prevented the executive from claiming that inherent authority over foreign investment is “vested” in the executive, even though investment authority is not specifically mentioned in the Constitution.

In addition, Congress’s actions in foreign investment increased the power previously exercised by the President acting alone. Before the passage of IEEPA and Exon-Florio, the President only monitored foreign investment in U.S. national security assets. The President had not claimed authority to block or amend transactions based on perceived national security threats. IEEPA and Exon-Florio, however, permit the President to actually block transactions that endanger national security. In FISA, by contrast, Congress set out a surveillance mechanism that sought to limit powers the President had previously exercised without Congressional input. By increasing presidential authority in foreign investment, rather than seeking to limit power already exercised by the executive, Congress is the source of the President’s authority, and therefore can more plausibly take back that authority in the future to serve as a greater check on the President’s power.

2. Judicial approval

Beginning with Exon-Florio and continuing with FINSA, foreign investment legislation has explicitly excluded judicial review of the President’s national security determinations. Because courts generally favor the executive in foreign relations decisions and at times have expanded the powers granted to the executive by Congress, this exclusion of judicial review probably also weighs in favor of maintaining separation of powers in foreign investment. Specifically, by excluding judicial review, FINSA and its precursors prevent courts from expanding executive powers beyond those granted by Congress or from legitimating a broad exercise of authority by the executive. To increase its own power here, Congress may consider excluding judicial review of all executive actions under FINSA except those that would constitute violations of the Constitution.
3. Congressional effectiveness

Further, unlike its actions (or inaction) with respect to international agreements, war, and surveillance, Congress has effectively responded to new threats to national security in foreign investment. When IEEPA proved ineffective because Presidents were reluctant to declare national emergencies to block particular transactions, Congress did away with this requirement in Exon-Florio. When foreign acquisition of military technology became a national security threat in the proposed LTV-Thomson transaction, Congress amended existing law to require consideration of technology concerns in CFIUS meetings. Finally, when energy became a national security concern after 9/11 and the CNOOC proposal, Congress required the consideration of energy concerns in the President’s national security considerations. By responding effectively in these situations, Congress has curtailed the need for unilateral executive action, therefore precluding the claims of inherent authority that have been made in other areas of foreign relations.

C. Policy recommendations

Though Congress has successfully avoided three major factors that have contributed to the expansion of executive power in other areas of foreign relations law, Congress could make several changes to FINSA in order to further secure its role in foreign investment. Specifically, the maintenance of separation of powers could be better preserved if future amendments to FINSA provided mechanisms for congressional comments on individual transactions, legislative participation in CFIUS, mandatory language in FINSA, and media scrutiny of executive action.

1. Individual comments

At present, FINSA and its precursors do not allow Congress to comment on individual transactions or set conditions on executive actions in foreign investment. In light of the successes of RUDs in international agreement and congressional limits on the use of force in particular conflicts, Congress should consider creating a mechanism that would facilitate legislative comments on individual decisions of the executive, indicating areas of approval and disapproval with respect to a particular transaction or action. Though such legislation would have to be carefully drafted to avoid an impermissible legislative veto, a mechanism that allowed Congress to comment on particular transactions before CFIUS had made its final decision would be feasible under current law.

2. Legislative participation

In addition to creating a mechanism for congressional comments on individual transactions, current investment laws could be amended to permit legislative participation or representation on CFIUS. Because the Committee already has two non-voting, ex officio members, some degree of participation by a legislator could probably be established without running afoul of the incompatibility clause.\textsuperscript{196} Even if a legislator cannot be included in the Committee as a full or an ex-officio member, Congress might choose an executive official to serve on the Committee with the specific task of representing the legislative interest in preserving separation of powers in foreign investment. As the operation of Frankfurter’s gloss theory in international agreements and war has demonstrated, the reporting requirements currently included in FINSA and its precursors are insufficient to preserve separation of powers.

3. Mandatory language

Though FINSA currently uses some mandatory language—requiring CFIUS to review transactions involving foreign governments and report its national security findings in all transactions to Congress—future amendments could strengthen this language. The absence of mandatory language is most obvious in FINSA’s broad definition of “national security.” By establishing a more concrete definition of national security, Congress could place more definite limits on the executive. Specifically, Congress could amend the current statute by establishing a hierarchy of mandatory factors for national security considerations and allowing permissive consideration of other factors. Under the current law, the executive enjoys nearly unfettered discretion in making national security determinations, which could lead to broader assertions of power by the executive.

In addition, Congress should add an exclusivity clause to FINSA, as included in FISA, stating that the CFIUS process is the sole constitutional means of reviewing and approving transactions that implicate U.S. national security interests. Such a clause would strengthen separation of powers by disputing any executive claims of inherent authority over foreign investment if the President seeks to claim that foreign investment constitutes an invasion.

4. Media scrutiny

Finally, as demonstrated above, the media has been a powerful check on presidential authority in other foreign relations matters. Though public scrutiny has already contributed to the frustration of several transactions considered national security threats, Congress could strengthen the role of the media by

\textsuperscript{196} See supra text accompanying note 78.
permitting the President’s “national security” determinations to be reported under FOIA. Even if immediate publication of these determinations is not feasible, Congress could provide a deadline at which such determinations would become public, such as ten years after the determination has been made. Knowledge that its decisions would be subject to media scrutiny at a future date may check the discretion of CFIUS in foreign investment and better protect separation of powers, while still preserving national security interests.

CONCLUSION

Under the framework established by applying Frankfurter’s gloss theory to other areas of foreign relations, Congress has been relatively successful in preserving separation of powers in foreign relations. Unlike its actions in international agreements, surveillance, and war, Congress has asserted its authority by legislating in foreign investment before the executive has asserted long-standing independent authority, by excluding judicial review of executive determinations under FINSA, and by effectively responding to new developments in the area.

Current legislation has several weaknesses, however, and Congress has not taken all possible steps to affirm its authority. In order to do so, Congress should amend FINSA to facilitate congressional comments on individual transactions in a process similar to the attachment of RUDs to the Senate’s consent to treaties. In addition, Congress should create a mechanism to incorporate a representative of legislative interests in CFIUS itself. Like FISA, legislation on foreign investment should be amended to include an exclusivity provision to preclude alternative actions by the executive to approve or block transactions involving national security assets. Finally, Congress should amend the statute to permit future FOIA reviews so that the prospect of future media analysis of CFIUS reviews could curb executive discretion.
**FREE ENTERPRISE FUND v. PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD: THE DECISION THAT CORPORATE AMERICA MAY FOREVER BE WAITING FOR**

Chad N. Eckhardt*

I. INTRODUCTION

The day Senator Paul Sarbanes (D-Md.) introduced his bill¹ to the Senate, WorldCom publicly announced that it had overstated earnings during the past five quarters by $3.8 billion.² Needless to say, Senator Sarbanes had the situational backing his bill needed. Within three weeks, Senator Sarbanes’s bill was unanimously passed in the Senate.³ Two weeks thereafter, the bill had been reconciled with a similar House Bill, presented to both houses of Congress, and passed⁴ as the Sarbanes-Oxley Act of 2002.⁵

This highly reformatory Act prompted speculation about its constitutionality.⁶ This speculative debate was given a concrete foundation when the Free Enterprise Fund of Washington D.C. and Beckstead and Watts, LLP, challenged the constitutionality of the Act.⁷ *Free Enterprise Fund v. Public Company Accounting Oversight Board* is the only case challenging whether the Public Company Accounting Oversight Board (hereinafter the “Board”) violates the Appointments Clause of the United States Constitution or the separation of powers doctrine.

This note focuses on whether the District of Columbia Circuit Court of Appeals’ decision that the Board does not violate the Appointments Clause and the separation of powers doctrine is correct. Section II reviews the brief history

---

¹ Chad N. Eckhardt is a J.D. candidate for 2009 at the Salmon P. Chase College of Law, Northern Kentucky University. He earned his M.A. in Education from the University of Kentucky in 2000 and his B.A. in English from the University of Kentucky in 1999.


³ Donna M. Nagy, *Playing Peekaboo with Constitutional Law: The PCAOB and Its Public/Private Status*, 80 NOTRE DAME L. REV. 975, 1005 (2005) (“[O]n June 25, 2002, after WorldCom’s shocking revelation that it had overstated its earnings by more than $3.8 billion during the past five quarters, primarily because it had improperly accounted for its operating costs. Senator Sarbanes introduced Senate Bill 2673 to the full Senate that very same day . . . .” (citing Robert Frank et al., *Scandal Scoreboard*, WALL ST. J., Oct. 3, 2003, at B1)).

⁴ Id.

⁵ Id. at 1005-06.


⁷ Nagy, supra note 2, at 979. Nagy’s article is the first to consider the constitutionality of the Public Co. Accounting Co. Oversight Bd.

of the Sarbanes-Oxley Act, as well as the Appointments Clause and separation of powers jurisprudence. Section III provides the factual and procedural background culminating in the Circuit Court’s decision as well as a detailed summary of the judges’ opinions. Section IV analyzes whether the Circuit Court’s decision is correct based on Supreme Court precedent. Finally, Section V concludes the note.

II. BACKGROUND

President George W. Bush heralded The Sarbanes-Oxley Act as the “most far-reaching reforms of American business practices since the time of Franklin Delano Roosevelt.” Enacted on July 30, 2002, in the wake of serious corporate scandals, such as WorldCom, Enron, and Tyco, the Sarbanes-Oxley Act of 2002 attempted to restore national confidence in the nation’s securities. Given this intense public and political concern, it is not surprising the Act passed nearly without opposition.

Leaders in the corporate area did not offer such a showing of approval. Various portions of the Act required major changes to corporate compliance and auditing, thus, creating substantial increases in costs. One study determined the aggregate cost to the economy to be well over one trillion dollars.

---

9. See Lawrence A. Cunningham, The Sarbanes-Oxley Yawn: Heavy Rhetoric, Light Reform (And It Just Might Work), 35 CONN. L. REV. 915, 924-25 (2003); see also Terry Morehead Dworkin, SOX and Whistleblowing, 105 MICH. L. REV. 1757, 1758 (2007) (“It was the wrongdoing, scandals, and resultant publicity, and the anger brought on by the leaders of failed corporations such as Enron and WorldCom that were the particular impetus for the law.”). For a list of all the companies that suffered serious financial scandals, see Perry E. Wallace, Accounting, Auditing Audit Committees After Enron, et al.: Governing Outside the Box Without Stepping Off the Edge in the Modern Economy, 43 WASHBURN L.J. 91, 102-03 nn.56-64 (2003).
11. Nagy, supra note 2, at 1006 (noting that the Sarbanes-Oxley Act passed the House with a vote of 423-3 and the Senate with a vote of 99-0).
14. Kenneth W. Starr, A Verdict on Sarbanes-Oxley: Unconstitutional, WALL ST. J., Dec. 16-17, 2006, at A11 (“Sarbox has cost the U.S. economy over $1 trillion, according to one study published by the AEI-Brookings Center.”).
In order to provide greater oversight of public companies’ accounting practices, the Act enhanced federal oversight of the accounting industry through the creation of the Board. Congress created the Board to oversee the accounting industry that many thought was enabling major companies, during their downward spirals, at the expense of public investors. Thus, the Board was given immense responsibility. The Board is authorized to “register public accounting firms.” Additionally, the Board is required to establish “auditing, quality control, ethics, independence, and other standards relating to the preparation of audit reports” for public accounting firms issuing such reports. Furthermore, the Board is empowered to “conduct inspections of registered public accounting firms” either annually or every three years depending on whether the public accounting firm creates audit reports for more than 100 issuers per year. The Board is also empowered to “conduct investigations and disciplinary proceedings concerning, and impose appropriate sanctions where justified upon, registered public accounting firms.” Specifically, the Board has the duty to “enforce compliance with this Act, the rules of the Board, professional standards, and the securities laws relating to the preparation and issuance of audit reports.” Although many of these express powers demonstrate a new Congressional approach to handling private matters, the manner in which Congress ensured the Board’s independence was extremely unique.

The most novel aspect of this legislation is the status of the Board. “The Board shall not be an agency or establishment of the United States Government, and . . . shall be subject to, and have all the powers conferred upon a nonprofit corporation.” This quasi-private status furthers the “deliberate legislative choice to create an independent entity.” Given the broad duties conferred and the independent status bestowed upon the Board, the Board appeared to be susceptible to constitutional challenge.

The most susceptible aspect of the Board appears to be the manner in which board members are appointed. “[T]he [Security Exchange Commission], after consultation with the Chairman of the Board of Governors of the Federal Reserve System and the Secretary of Treasury, shall appoint the chairperson and other initial members of the Board.” Vacancies are to be “filled in the same

16. Id. § 7211(c)(1).
17. Id. § 7211(c)(2).
18. Id. § 7211(c)(3).
19. Id. § 7214(b)(1).
20. Id. § 7211(c)(4).
22. Id. § 7211(b). For a thoughtful discussion concerning the Board’s status, see Nagy, supra note 2, at 1006-26.
manner as provided for appointments under this section." This scheme provides for the appointment of the Board members without any action by the President. Given the broad duties conferred on the Board and the lack of Presidential participation in the appointment of the Board members, speculation concerning whether Congress’ scheme creating the Board violates the Appointments Clause or the separation of powers doctrine surfaced. A brief overview of each principle provides the necessary principles upon which the Circuit Court’s opinions rest.

A. Appointments Clause

The Appointments Clause of Article II of the United States Constitution reads as follows:

[The President] shall nominate, and by and with the Advice and Consent of the Senate, shall appoint Ambassadors, other public Ministers and Consuls, Judges of the supreme Court, and all other Officers of the United States, whose Appointments are not herein otherwise provided for, and which shall be established by Law: but Congress may by Law vest the Appointment of such inferior Officers, as they think proper, in the President alone, in the Courts of Law, or in the Heads of Departments.

So few cases intricately involving this clause have arisen that the District of Columbia Circuit Court stated in 1988, “[t]wo hundred years after the adoption of the United States Constitution the federal courts are, essentially for the first time, required to construe closely the appointments clause of Article II.” Despite this relatively new inquiry, a longstanding principle is that this clause “for purposes of appointment . . . divides all its officers into two classes,” principal officers and inferior officers. “Principal officers are selected by the President with the advice and consent of the Senate. Inferior officers Congress may allow to be appointed by the President alone, by the heads of the departments, or by the Judiciary.” In an inquiry involving the Appointment Clause, “[t]he initial question is, accordingly, whether [the officer] is an ‘inferior’ or a ‘principal’ officer.”

Although the question is simple, the determination is not. “The line between ‘inferior’ and ‘principal’ officers is one that is far from clear, and the Framers

25. Id. § 7211(e)(4)(B).
27. U.S. CONST. art. II, § 2, cl. 2.
provided little guidance into where it should be drawn."32 In the first substantial Appointment Clause case, the Court declined to decide where the line should be drawn,33 but it articulated four factors which have since begun the inquiry: 1) whether the appointee is subject to removal by a higher Executive officer, 2) the extent to which the appointee’s duties are limited, 3) the extent to which the appointee’s jurisdiction is limited, and 4) the extent to which the appointee’s tenure is limited.34 Because “Morrison did not purport to set forth a definitive test,”35 the Supreme Court in Edmond v. United States felt confident in declaring “it evident that ‘inferior officers’ are officers whose work is directed and supervised at some level by others who were appointed by Presidential nomination with the advice and consent of the Senate.”36 This “evident” statement in Edmond has not merely relegated Morrison to the laundry list of cases making such a determination.37 Rather, courts continue to rely on the four factors articulated in Morrison to delineate a line that is “far from clear” in most situations.38 Thus, while the test articulated by Justice Scalia in Edmond is the most recent Supreme Court espousal on the issue, consideration of the four factors that Morrison found determinative helps analyze what “directed and supervised” means.

B. Separation of Powers

Although the idea that all three branches of the government must be separate is not articulated in the Constitution,39 such “is implied in the very fact of the separation of the powers of these departments by the Constitution.”40 Furthermore, “[t]he fundamental necessity of maintaining each of the three

---

32. Id. at 671.
33. Id. at 671 (“We need not attempt here to decide exactly where the line falls between the two types of officers . . . .”).
34. Id. at 671-72. These factors were later applied in Edmond v. United States, 520 U.S. 651, 661 (1997).
35. Edmond, 520 U.S. at 661.
36. Id. at 663.
37. Opinions routinely include a paragraph rapidly listing the precedents that have found officers to be inferior or principal. See Edmond, 520 U.S. at 661; United States v. Vazquez, 69 F. Supp. 2d 286, 289 (D.P.R. 1999).
38. Stanley v. Gonzales, 476 F.3d 653, 659 (9th Cir. 2007) (“Under Morrison, we examine several factors to determine whether the officer is a principal or an inferior officer. First, we inquire whether the officer is subject to removal by a higher Executive Branch official. Second, we look at whether the officer is empowered to perform only certain, limited duties. Finally, we look at whether an officer's duties are limited in jurisdiction. Generally, if an officer is limited in ‘tenure, duration, [and] . . . duties,’ those findings lead to the conclusion that she is an inferior officer.” (internal citations omitted)).
39. Free Enterprise Fund v. Public Co. Accounting Oversight Bd., 537 F.3d 667, 678 (D.C. Cir. 2008) (“Although not expressly included in the Constitution itself, the principle of separation of powers is implicit in the first three articles of the Constitution that define separate roles for the legislative, executive, and judicial branches.” (internal citations omitted)).
general departments of government entirely free from the control or coercive influence, direct or indirect, of either of the others, has often been stressed and is hardly open to serious question.\textsuperscript{41} These basic propositions underlie the touchstone case\textsuperscript{42} for resolving whether restrictions to the President’s power to remove an officer violate the separation of powers doctrine. \textit{Humphrey’s Executor v. United States} upheld the for-cause removal restriction placed upon the President’s ability to remove a Federal Trade Commissioner.\textsuperscript{43}

Over fifty years later, the Supreme Court added to this rather narrow field of jurisprudence. In \textit{Morrison v. Olson}, the Court stressed that “[t]he analysis contained in our removal cases is . . . to ensure that Congress does not interfere with the President’s exercise of the ‘executive power’ and his constitutionally appointed duty to ‘take care that the laws be faithfully executed’ under Article II [of the Constitution].”\textsuperscript{44} Because this determination “will depend upon the character of the office”\textsuperscript{45} at issue, the answer will inevitably be made on a case-by-case basis.\textsuperscript{46} However, \textit{Morrison} stated that “the real question is whether the removal restrictions are of such a nature that they impede the President’s ability to perform his constitutional duty.”\textsuperscript{47}

III. FACTS OF FREE ENTERPRISE FUND V. PUBLIC COMPANY 
ACCOUNTING OVERSIGHT BOARD

A. The Complaint

On February 7, 2006 the Free Enterprise Fund of Washington D.C., a non-profit organization,\textsuperscript{48} along with Beckstead and Watts, LLP, a Nevada accounting firm,\textsuperscript{49} filed a complaint in the United States District Court for the District of Columbia alleging that the Sarbanes-Oxley Act of 2002 violated the United States Constitution.\textsuperscript{50} First, the complaint alleged that the Board violated

\begin{itemize}
  \item \textsuperscript{41} Id. at 629.
  \item \textsuperscript{42} Id. As of this writing, \textit{Humphrey’s Executor} has 324 citing decisions and 694 law review citations.
  \item \textsuperscript{43} \textit{Humphrey’s Executor}, 295 U.S. at 631-32.
  \item \textsuperscript{44} \textit{Morrison v. Olson}, 487 U.S. 654, 689–90 (1988) (referencing the President’s responsibility to “take Care that the Laws are faithfully executed.” (citing U.S. CONST. art. II, § 3)).
  \item \textsuperscript{45} \textit{Humphrey’s Executor}, 295 U.S. at 631, cited with approval in \textit{Morrison}, 487 U.S. at 687.
  \item \textsuperscript{46} Humphrey’s Executor v. United States, 295 U.S. 602, 632 (1935) (“To the extent that, between the decision in the Myers Case, which sustains the unrestricted power of the President to remove purely executive officers, and our present decision that such power does not extend to an office such as that here involved, there shall remain a field of doubt, we leave such cases as may fall within it for future consideration and determination as they may arise.”).
  \item \textsuperscript{47} \textit{Morrison}, 487 U.S. at 691.
  \item \textsuperscript{49} Id.
  \item \textsuperscript{50} Complaint, supra note 7, at ¶ 1.
\end{itemize}
the Appointments Clause. The complaint stated that “members of the Board are principal officers whose appointments must be made by the President by and with the advice and consent of the Senate.” Alternatively, if “the members of the Board are inferior officers[,]” according to the complaint, “the appointment of Board members by the SEC violate[d] the Appointments Clause” because the “SEC is not a department within the meaning of the Clause.”

Second, the plaintiffs argued that the Board violated the separation of powers doctrine. The complaint alleged that “[t]he Board’s exercise of wide-ranging, core executive power, immune from Presidential oversight, impermissibly impedes and undermines the President’s ability to perform his constitutional duties and prerogatives.”

Finally, although not the subject of this note, the complaint alleged unconstitutional delegation in that “the Act improperly and unconstitutionally delegate[d] legislative power to an entity outside the Legislative Branch.”

B. The District Court’s Opinion

On March 21, 2007, the United States District Court for the District of Columbia granted summary judgment in favor of the defendants. In granting the defendants’ motion for summary judgment, Judge James Robertson found that the Appointments Clause was not violated because “the PCAOB members are inferior officers.” Although Judge Robertson succinctly addressed the challenge, he found that Board members fit squarely into the precedent. On the alternative Appointments Clause argument, Judge Robertson found that the plaintiffs lacked standing.

The other arguments were equally unavailing. Judge Robertson quickly dismissed the separation of powers argument when he declared that “[t]he Supreme Court has never held that the Constitution requires the President to maintain direct removal power over inferior officers.” Finally, Judge Robertson disregarded the overreaching delegation when he found the standards

51. Id. at ¶¶ 86-92.
52. Id. at ¶ 90.
53. Id., at ¶ 91.
54. Id. at ¶¶ 81-85.
55. Id. at ¶ 85.
56. Complaint, supra note 7, at ¶ 95.
58. Id. at *12.
59. Id. (finding Board members to be similar to judges of the Coast Guard Court of Criminal Appeals as in Edmond v. United States, 520 U.S. 651 (1997), administrative law judges as in Landry v. FDIC, 204 F.3d 1125 (D.C. Cir. 2000), and independent counsel as in In re Sealed Case, 829 F.2d 50 (D.C. Cir. 1987)).
60. Id. at *14 (“On this point, however – that PCAOB members should have been appointed by the SEC Chairman rather than by the entire Commission – the plaintiffs lack standing.”).
61. Id. at *15.
that the Board were to apply were “intelligible . . . [and] squarely within the bounds of modern non-delegation doctrine.”

Notably, Judge Robertson stated that the plaintiffs have presented “nothing but an hypothetical scenario of an over-zealous or rogue PCAOB investigator.”

The plaintiffs appealed Judge Robertson’s decision to the United States Court of Appeals for the District of Columbia. After the oral argument on April 15, 2008, corporate America eagerly awaited a response. Finally, after a long summer, the United States Court of Appeals for the District of Columbia issued its opinion on August 22, 2008. In a split decision, the court affirmed Judge Robertson’s ruling.

C. The Circuit Court’s Opinion

The court split on both issues. As to the challenged violation of the Appointment Clause, the majority stated that the test of “[w]hether one is an ‘inferior’ officer [under Edmond] depends on whether he has a superior.” The majority held that “the Act does not encroach upon the Appointment power because, in view of the Commission’s comprehensive control of the Board, Board members are subject to direction and supervision of the Commission and thus are inferior officers not required to be appointed by the President.” The majority cited numerous applications of the “Commission’s comprehensive control of the board,” such as the Commission’s approval of Board rules, de novo review of board adjudications, and ability to alter board sanctions. Additionally, the majority found that “[j]ust as in Morrison, ‘the fact that [Board members] can be removed by the [Commission] indicates that [they are] to some degree ‘inferior’ in rank and authority.”

In the dissent however, Judge Kavanaugh framed the Edmond analysis differently. First, Edmond would require the SEC “to prevent and affirmatively

---

62. Id. at *17.
65. Id.
66. Id.
67. Id.
68. Id. at 672 (citing Edmond v. United States, 520 U.S. 651, 662 (1997)).
69. Id. at 669.
71. Id. at 670 (“All Board adjudications are subject to the Commission’s de novo review.” (citing 15 U.S.C. § 7217(c)(2) (2008))).
72. Id. (“The Commission is empowered to ‘enhance, modify, cancel, reduce, or require the remission of a sanction imposed by the Board.’” (citing 15 U.S.C. § 7217(c)(3) (2008))).
73. Id. at 674 (citing Morrison v. Olson, 487 U.S. 654, 671 (1988)).
command, and to manage the ongoing conduct of, Board inspections, Board investigations, and Board enforcement actions.” Judge Kavanaugh did, however, recognize the Commission’s control over the Board, but he did not characterize the control as comprehensive; rather, he considered the measures “cobble[d] together disparate pieces of statutory text.” Specifically, Judge Kavanaugh highlighted the fact that approval of Board rules only existed before they take effect and such review has no relation to the Board’s ability to conduct inspections, investigations, and enforcement actions. As such, the direction of the SEC over the Board is not comprehensive. Additionally, Judge Kavanaugh found the review of Board sanctions to be an “[a]fter-the-fact” and “quasi-judicial review” that “does not remotely equate to direction and supervision for purposes of Edmond.” His poignant analogy is noteworthy, “[o]ne would not say, for example, that a U.S. Attorney is directed and supervised by a federal district court in his or her investigative decisions just because a court ultimately would have an opportunity to review any indictment or subpoena challenge.”

Finally, the dissent found the limitations on the Board’s investigation and enforcement responsibilities by the SEC to be a “theoretical power to alter the Board’s jurisdiction [that] does not equate to power to prevent and affirmatively command, and to manage the ongoing conduct of, Board inspections[.]” Given the difference in these textual considerations, Judge Kavanaugh found the members of the Board to be principal officers.

Notably, Judge Kavanaugh relied heavily upon the fact that Board members are only removable for cause by the Commission. Since “[r]emovability at will carries with it the inherent power to direct and supervise[,]” Judge Kavanaugh stated, “the key initial question in determining whether an executive officer is inferior is whether the officer is removable at will.” Because for cause removal provides “substantive independence from direction and supervision,” Judge Kavanaugh reasoned, “for cause officers ordinarily are not ‘directed and

---

74. Id. at 687 (Kavanaugh, J., dissenting).
75. Id. at 711 (Kavanaugh, J., dissenting).
76. Free Enterprise Fund v. Public Co. Accounting Oversight Bd., 537 F.3d 667, 710 (D.C. Cir. 2008) (Kavanaugh, J., dissenting) (“[T]he Board argues that the SEC has power to review Board rules before they take effect. That is true . . . [S]ection 7217 in turn gives the SEC power to review only Board rules before they take effect.”).
77. Id. (“To qualify as inferior, an officer must be statutorily subject to direction and supervision in all significant activities.”) (emphasis added)).
78. Id.
79. Id.
80. Id.
81. Id. at 709.
84. Id. (citing Edmond, 520 U.S. at 663).
supervised’ for purposes of Edmond” and “they presumptively should be considered principal officers.” 85

With respect to the separation of powers argument, the majority fully embraced the doctrine but highlighted that “this principle, however, ‘by no means contemplates total separation of each of these three essential branches of Government.’” 86 The majority compared the restrictions on presidential power here with those found constitutional in Morrison. In Morrison, the majority noted, “neither the President nor the Attorney General had the power to appoint the Independent Counsel or the power to control her investigatory or prosecutorial authority.” 87 Additionally, “the Independent Counsel possessed significant independence from the President.” 88 Here, on the other hand, “the President possesses significant influence over the Commission, which in turn possesses comprehensive control over the Board.” 89 The result of such a comparison, according to the majority, is that “the President’s powers under the Act extend comfortably beyond the minimum required to ‘perform his constitutionally assigned duties.’” 90

Despite the contention “that the double for-cause limitation on removal makes it impossible for the President to perform his duties[,]” the majority considered this removal power “one of several criteria relevant” to the determination. 91 Additionally, the majority stated that “the Fund bears a heavy burden” because it presents a “facial challenge.” 92 In sum, “the Act as a whole provides ample Executive control over the Board.” 93

Rather than recognize that the separation of powers doctrine does not contemplate total separation as the majority, Judge Kavanaugh’s focus began solely with intrusions into the executive’s removal power. Judge Kavanaugh began his analysis with Myers, “the Myers Court made clear that Congress could play no role in the removal of executive officers.” 94 As such, “[t]he PCAOB

85. Id.
86. Id. at 679 (citing Buckley v. Valeo, 424 U.S. 1, 121 (1976)).
87. Id. at 681 (citing Morrison v. Olson, 487 U.S. 654 (1988)).
88. Id.
90. Id. (citing Morrison, 487 U.S. at 696).
91. Id. at 679.
92. Id. at 684 n.14 (“Contrary to our dissenting colleague, Dis. Op. at 704 n. 12, the fact that this is a facial challenge significantly affects the analysis, for the Fund bears a heavy burden to demonstrate that the Act unduly constrains the President’s ability to see that the laws are faithfully executed in all circumstances and cannot be constitutionally applied . . . .”).
93. Id. at 683.
94. Id. at 694 (Kavanaugh, J., dissenting). Judge Kavanaugh further stated, “[i]n its 1986 decision in Bowsher v. Synar, for example, the Court reaffirmed this precise holding.” Id. The strength of this proposition is severely challenged by Chief Justice Rehnquist in Morrison when he notes that the argument that the President must have absolute discretion to discharge “purely executive” officials at will “was raised by the Solicitor General in Bowsher v. Synar, although as Justice White noted in dissent in that case, the argument was clearly not accepted by the Court at
would be flatly unconstitutional.”

However, Judge Kavanaugh acknowledged that “Humphrey’s Executor and Morrison authorize a significant intrusion on the President’s Article II authority.” Judge Kavanaugh’s depiction of both cases allows only for-cause restrictions on the President or Presidents alter ego’s removal power. Since Judge Kavanaugh viewed Humphrey’s Executor and Morrison as “the outermost constitutional limits” of removal restrictions, his analysis depends upon whether the Board “further attenuate[s] the President’s control over executive officers” than either scheme present in Humphrey’s Executor or Morrison.

Judge Kavanaugh emphasized that the President may only remove a Commissioner of the SEC for cause who, in turn, may only remove a Board member who “has ‘willfully’ broken the law [or] ‘willfully abused’ his or her authority.” As such, under the congressional scheme here, “the President is two levels of for-cause removal away from [the Board].” Thus, “the Sarbanes-Oxley Act completely strip[s] the President’s ability to remove [Board] members” and “eliminate[s] any meaningful Presidential control over the [Board].”

Ultimately, Judge Kavanaugh found the removal restrictions to violate the separation of powers doctrine for a variety of reasons. First, Judge Kavanaugh thought the court should “hold the line and not allow encroachments on the President’s removal power beyond what Humphrey’s Executor and Morrison already permit.” Second, he found that language in Morrison “all but resolves the removal issue in this case.” Third, “the most telling indication of the

---

95. Free Enterprise Fund v. Public Co. Accounting Oversight Bd., 537 F.3d 667, 694 (D.C. Cir. 2008) (Kavanaugh, J., dissenting) (“The PCAOB would be flatly unconstitutional because the statute restricts the President’s power to remove PCAOB members at will . . . .”).

96. Id. at 696.

97. Id. (“[B]ecause the President’s alter ego (the Attorney General) retained the authority to remove the independent counsel for cause, the President’s ‘power to remove’ was not ‘completely stripped.’” (quoting Morrison, 487 U.S. at 692)).

98. Id. at 698.

99. Id. at 697-98 (“The question is whether the Sarbanes-Oxley Act’s restriction on the President’s removal power over the PCAOB is unconstitutional under Humphrey’s Executor and Morrison.”).

100. Id. at 697 n.7 (citing 44 U.S.C. § 3502(5) (2000)).


102. Id. at 686.

103. Id. at 697.

104. Id. at 698.

105. Id. (referencing the following language from Morrison, “This is not a case in which the power to remove an executive official has been completely stripped from the President, thus providing no means for the President to ensure the ‘faithful execution’ of the laws. Rather, because the independent counsel may be terminated for ‘good cause,’ the Executive, through the Attorney General, retains ample authority to assure that the counsel is competently performing his or her statutory responsibilities . . . .” 487 U.S. 654, 692 (1988)).
severe constitutional problem with the [Board] is the lack of historical precedent for this entity."106 Finally, upholding the structure here would enable “Congress to create a host of similar entities . . . thereby splinter[ing] executive power to a degree not previously permitted.”107

IV. ANALYSIS

The manner in which Board members are appointed does not violate the Appointments Clause because the members are inferior officers. The removal restrictions, however, violate the separation of powers doctrine with respect to the President’s removal power. Despite the limited precedent available, the President must retain more removal power over officers who perform executive functions than the removal scheme created by the Sarbanes-Oxley Act allows.

A. The Appointments Clause

The initial analysis in any Appointments Clause challenge concerns whether the appointment being challenged is that of an officer or merely an employee.108 Here, there is little doubt that the Board members are officers despite the public/private status that Congress envisioned for the Board.109 If the position is that of an officer, then the next determination is whether the officer is an inferior or principal one.110 Thus, to determine whether the Appointments Clause requires the President to appoint an officer “with the advice and consent of the Senate,” one must apply the test articulated by Justice Scalia in Edmond. However, if one overlooks the analysis of Morrison, the case from which Justice Scalia, in writing for the majority, distilled his test, a major analytical link is ignored. Thus, whether members of the Board must be appointed by the President with the advice and consent of the Senate, one must apply the test articulated by Justice Scalia in Edmond. However, if one overlooks the analysis of Morrison, the case from which Justice Scalia, in writing for the majority, distilled his test, a major analytical link is ignored. Thus, whether members of the Board must be appointed by the President with the advice and consent of the Senate depends upon whether the officers’ “work is directed and supervised at some level by others who were appointed by presidential nomination with the advice and consent of the

106. Id. at 699.
110. A final determination, in some instances, involves whether the officer was appointed by the “heads of the departments” as required by the Appointments Clause. This case presents such a situation; however, resolution of this issue is beyond the scope of this note. See infra note 135.
Because Justice Scalia’s test is taken from Chief Justice Rehnquist’s opinion in *Morrison*, the factors that led the majority in *Morrison* to find that the position of an Independent Counsel “clearly falls on the ‘inferior officer’ side of th[e] line” illuminate the meaning of Justice Scalia’s phrase “directed and supervised at some level[].”

The members of the Board are only able to be removed by the Securities Exchange Commission (SEC) if the member has “willfully broken the law [or] willfully abused his or her authority.” As the Commissioners of the SEC are “appointed by presidential nomination with the advice and consent of the Senate,” the Board member is “to some degree ‘inferior’ in rank and authority.”

However, as Judge Kavanaugh suggested, the for-cause removal here is more limited than that previously considered. Normally, the for-cause removal includes “language of inefficiency, neglect of duty, or malfeasance.” Here, a Board member may only be removed if he or she “willfully violated any provision of th[e] Act, the rules of the Board, or the securities laws” or “willfully abused th[e] authority of the member.” This language clearly limits removal more than the traditional language as Judge Kavanaugh suggests. Yet, with respect to enforcement of the Board’s rules or professional standards for accountants, the standard is merely “without reasonable justification or excuse.” The removal of a Board member who fails to enforce compliance “without reasonable justification or excuse” is more akin to “inefficiency, neglect of duty, or malfeasance” than it is to “willful” conduct. Thus, although Board members are only removable for willful behavior in some circumstances, with respect to the quintessential executive function of enforcement, they are subject to the traditional removal standard.

Second, while the Board’s duties may be extensive with respect to the accounting industry, they are nevertheless limited in a manner consistent with *Morrison* and *Edmond*. The Independent Counsel in *Morrison* was granted “full

---

112. See supra Part II A.
114. Edmond, 520 U.S. at 663.
116. Id. at 672 (citing Edmond, 520 U.S. at 663) (“The Commissioners, who serve staggered five-year terms, are ‘appointed by Presidential nomination with the advice and consent of the Senate.’”).
118. Free Enterprise Fund, 537 F.3d. at 703 (Kavanaugh, J., dissenting) (“This is more restrictive removal language than the traditional for-cause language of inefficiency, neglect of duty, or malfeasance.”).
119. Id.
121. Id. § 7217(d)(3)(C).
power and independent authority to exercise all investigative and prosecutorial functions and powers of the Department of Justice.”122 Despite these extensive “investigative and prosecutorial functions,” the Court reasoned that the Independent Counsel was, nevertheless, “empowered by the Act to perform only certain, limited duties.”123 Further, the Independent Counsel was not authorized “to formulate policy” or perform “administrative duties outside of those necessary to operate her office.”124 Similarly, the Board does not “formulate policy” or perform unnecessary administrative duties. The Board is only authorized to establish standards for the accounting industry “subject to action by the Commission.”125 Additionally, the functions that the Board performs are then only in furtherance of compliance with those standards, which is necessary to the operation of the Board.126 Despite the significant role that the Board performs with respect to the accounting industry, Edmonds emphasized that the exercise of significant authority marks not the line between principal and inferior officer, but the line between officer and non-officer.127

Finally, the Board’s office, however, is not limited in jurisdiction or tenure in a fashion similar to the Independent Counsel in Morrison. Although the Board is empowered to audit and inspect only certain registered firms,128 such a limitation encompasses much more than that which was contemplated in Morrison. In Morrison, the Court found the Independent Counsel’s jurisdiction to be limited because she could only investigate “certain federal officials suspected of certain serious federal crimes” and then only within the scope granted by the Special Division.129 Likewise, the Board’s tenure is not limited “to accomplish a single task [at the end of which] the office is terminated”130 as was the case in Morrison. However, the Court in Edmond found that military judges are not limited in jurisdiction or tenure like the Independent Counsel in Morrison, but, did not find that dispositive.131

Even though the Board is not limited in jurisdiction or tenure, they are removable for unreasonably failing to enforce the Act and are subject to “explicit

---

123. Id. at 671.
124. Id. at 671-72.
126. Id. § 7211(c)(5) (stating that the Board may perform such other duties or functions necessary or appropriate to carry out the Act).
127. Edmond v. United States, 520 U.S. 651, 662 (1997) (“The exercise of ‘significant authority pursuant to the laws of the United States’ marks, not the line between principal and inferior officer for Appointments Clause purposes, but rather, as we said in Buckley, the line between officer and nonofficer.” (citation omitted)).
130. Edmond, 520 U.S. at 661 (quoting Morrison, 487 U.S. at 672).
131. Id. (“[T]he last two of these conclusions do not hold with regard to the office of military judge . . . . However, Morrison did not purport to set forth a definitive test . . . .”)
and comprehensive oversight by the Commission. Although Judge Kavanaugh questioned the practical effect of the Commission’s oversight, Edmond only requires that the work be “directed and supervised at some level,” not specifically on a level that enables the Commission “to manage the ongoing conduct of, Board inspections, Board investigations, and Board enforcement actions.” Thus, the majority is correct in determining that the Board members are inferior officers, and, therefore, their appointment is not in violation of the Appointments Clause.

With respect to Judge Kavanaugh’s dissent, it is important to discuss the emphasis he places on for-cause removal when he stated that for-cause officers “presumptively should be considered principal officers.” Judge Kavanaugh’s treatment of for-cause officers is understandable, yet ingenious. Essentially, Judge Kavanaugh took one of the factors from Morrison and converted it into a presumption while offering no authority for such treatment and none can be inferred. The fact that for-cause officers have “substantive independence” was not overlooked in Morrison. To the contrary, Chief Justice Rehnquist stated in Morrison, “insofar as [appellant] possesses a degree of independent discretion . . . the fact that she can be removed by the Attorney General indicates that she is to some degree ‘inferior’ in rank and authority.” Chief Justice Rehnquist’s statement implies that the removability factor is more important than independence factor because the ability to be removed indicates inferiority. Justice Scalia recognized this point in Edmond, which Judge Kavanaugh

134. Free Enterprise Fund, 537 F.3d at 687 (Kavanaugh, J., dissenting).
135. Id. The alternative argument presented by the plaintiffs, that the appointment of Board members by the SEC and not the Chairperson of the SEC is a violation of the Appointments Clause because the Commissioners are not the head of the Commission, is beyond the scope of this note. However, the majority found that “[a]ll three branches of government are in agreement that the head of an agency can be a multi-member body.” Id. at 677. Additionally the majority relied upon Silver v. United States Postal Service, 951 F.2d 1033 (9th Cir. 1991), which found that “the [nine governors] are the head of the department.” Id. at 1038. Although Judge Kavanaugh did not need to address the issue because he found the primary Appointments Clause argument persuasive, Judge Robertson, in the district court decision, stated that the contention by the PCAOB “that the ‘head’ of the SEC is the SEC as a whole . . . is a stretch.” Free Enterprise Fund v. Public Co. Accounting Oversight Bd., No. 06CV0217, 2007 U.S. Dist. LEXIS 24310, at *13 (D.D.C. Mar. 21, 2007). Ultimately, Judge Robertson found that “[m]ulti-member bodies may, on occasion, properly constitute heads of departments for Appointments Clause purposes, but the SEC is not of one of them.” Id. at *14.
137. Id. at 708 n.16 (“Presuming for-cause officers to be principal officers makes great sense” because the inherent independence of for-cause removal dictates “full scrutiny of the officers’ character and qualifications” on the front-end. (citing Laurence H. Tribe, 1 AMERICAN CONSTITUTIONAL LAW § 4-8, at 684 (3d ed. 2000))).
138. Id. at 707.
referenced in the quoted passage, when he stressed the hierarchical structure of an officer having a superior, not the manner in which the officer could be removed. Despite Justice Scalia and Chief Justice Rehnquist’s language, Judge Kavanaugh created a presumption that for-cause officers should be considered principal. This presumption could misguide the analysis in a case like this one which is not “a relatively easy case in which to apply the ‘directed and supervised’ test.”

B. Separation of Powers

As stated above, the test determining whether the restrictions on the President’s removal power violate the doctrine of separation of powers is articulated in Morrison. Whether the removal restrictions violate the doctrine of separation of powers depends upon “whether the removal restrictions are of such a nature that they impede the President’s ability to perform his constitutional duty.” The constitutional duty to which Chief Justice Rehnquist referred is the duty to “take Care that the Laws be faithfully executed.”

The Board is authorized by Congress to promulgate rules that establish “standards relating to the preparation of audit reports.” Although this delegation is legislative in nature, the Board is also empowered to “enforce compliance with this Act.” Further, the Board shall “assess the degree of compliance of each registered public accounting firm and associated person of that firm with this Act.” Enforcement of the laws is a quintessential executive function. Thus, the Board performs functions that embrace aspects of the legislative and executive branches. Additionally, the majority in Free Enterprise Fund expressly acknowledged that the Board’s powers were essentially executive, “as in the case of the Board, Congress identified the need for an entity within the Executive Branch.”

In Humphrey’s Executor, the Court found that the duties of the Federal Trade Commission (FTC) were neither “political nor executive, but predominately quasi-judicial and quasi-legislative.” Since the FTC did not exercise any executive power, the removal of any commissioner, which was

140. Free Enterprise Fund, 537 F.3d. at 707 (Kavanaugh J., dissenting) (“Edmond was a relatively easy case in which to apply the ‘directed and supervised’ test . . . .”).
141. Morrison, 487 U.S. at 691.
142. U.S. Const. art. II, § 3.
144. Id. § 7211(c)(6).
145. Id. § 7214(a).
147. Humphrey’s Executor v. United States, 295 U.S. 602, 628 (1935) (“To the extent that it exercises any executive function, as distinguished from executive power in the constitutional sense, it does so in the discharge and effectuation of its quasi-legislative and quasi-judicial powers . . . .”).
limited to “inefficiency, neglect of duty, or malfeasance in office[,]” did not interfere with the President’s ability to faithfully execute the law.

Two aspects of the Board distinguish this case from *Humphrey’s Executor*. First, the Board performs executive functions whereas the court determined that the FTC did not. This performance is essential to the extent of removal. In *Myers*, the Court held that restriction upon a postmaster’s removal was unconstitutional. Although the Supreme Court has since attempted to limit *Myers*, the case provides a powerful precedent with regard to executive officers. Despite the attacks, Myers stands for the proposition that the President’s power of removal shall be at its greatest with regard to purely executive officers. This idea is consistent with the ideas of our founding fathers, “[r]epresentative James Madison of Virginia and others argued that Article II [of the Constitution] provided the President plenary power to remove executive officers[.]” Thus, it appears that limitations upon the President’s power to remove officers are inversely proportional to the extent to which the officers perform executive functions. In other words, as the executive-type functions performed by the officers increase, the limitations on the President’s removal power decrease. Under this formula, the limitations upon the President’s power to remove Board members should be less than that upheld in *Humphrey’s Executor* where the office in question did not perform any executive function. However, this is not the case.

The second aspect distinguishing this situation from *Humphrey’s Executor* involves the restrictions upon the President’s power to remove Board members. The limitation of removal in *Humphrey’s Executor* was “inefficiency, neglect of duty, or malfeasance in office.” Board members, however, may only be removed for “willfully violat[ing] any provision of th[e] Act, the rules of the Board, or the securities laws[,] . . . willfully abus[ing] the authority of th[e] member,” or failing to enforce the Board’s rules or professional standards for accountants “without reasonable justification or excuse.” A willful standard is

150. *See Humphrey’s Executor* v. United States, 295 U.S. 602 (1935). In *Humphrey’s Executor*, the Court distinguished *Myers* on the basis that the FTC did not exercise any executive power. It stated that “[t]he office of a postmaster is so essentially unlike the office now involved that the decision in the Myers case cannot be accepted as controlling our decision here. A postmaster is an executive officer restricted to the performance of executive functions.” *Id.* at 627. Later, the Court in *Morrison v. Olson* questioned the validity of the argument that “the President must have absolute discretion to discharge ‘purely’ executive officials at will.” 487 U.S. 654, 688-89 (1988).
153. *Id.* at 620.
155. *Id.* § 7217(d)(3)(C).
clearly more restrictive than inefficiency, neglect, or malfeasance. However, the fact that Board members can be removed for failing to enforce rules or standards “without reasonable justification or excuse” is more consonant with the removal restrictions in *Humphrey’s Executor*. Unlike *Humphrey’s Executor*, however, the President here does not have the power to remove the officer for such an unjustified failure to enforce the rules—the Commission does. Thus, the President would need to remove enough Commissioners, who have for-cause removal only, to enable the Commission to find that the Board member acted in a manner that warrants removal. Such a scheme attenuates the President from exercising removal power over the Board, when such restriction should be at its least, to an extent more restrictive than that contemplated in *Humphrey’s Executor*.

Despite this anomaly, the majority in Free Enterprise Fund found that “the President’s powers under the Act extend comfortably beyond the minimum required.” The majority relied predominately upon the comparison of the removal of Board members to the removal of the Independent Counsel in *Morrison* without fully considering the removal scheme created by Sarbanes-Oxley. Such a comparison is misplaced. *Morrison* involved a situation that demanded more independence from the President than that at issue here. The Independent Counsel was empowered to investigate “certain federal officials suspected of certain serious federal crimes” including the “President and Vice President.” Potentially, the Independent Counsel could investigate the close associates of the President as well as the President himself. Such investigative powers would necessarily need to be independent of the President in order for the Independent Counsel to exercise any authority granted to her by Congress. The Board, however, does not need such independence from the President because the President is not a potential target of the Board’s exercise of power. The Board seeks to enforce rules regulating the accounting industry, which is an objective more closely aligned to the FTC than an independent counsel. Additionally in *Morrison*, as Judge Kavanaugh pointed out in his dissent, “the President through his alter ego (the Attorney General) still retained the authority to remove the independent counsel.” Thus, the majority’s finding that the “President’s powers under the Act extend comfortably beyond the minimum required[,]” is questionable.

The majority’s opinion concerning the separation of powers argument is further weakened by other considerations. Instead of reading *Myers* to presume

---

156. *Id.*
159. *Id.* at 661 n.2.
160. *Free Enterprise Fund*, 537 F.3d at 687 (Kavanaugh, J., dissenting).
161. *Id.* at 681.
fewer limitations on the President’s removal of officers exercising executive power, and hence a starting point against restriction, the majority started on the other side when Judge Rogers, writing for the majority, stated that “[f]acial challenges are disfavored.”\textsuperscript{162} Such a statement placed a heavy presumption against the Appellants and their argument that the restrictions limit the President’s power unconstitutionally. The majority continued, albeit in a footnote, “the fact that this is a facial challenge significantly affects the analysis, for the Fund bears a heavy burden to demonstrate that the Act unduly constrains the President’s ability to see that the laws are faithfully executed in all circumstances and cannot constitutionally be applied.”\textsuperscript{163} Although this proposition is generally accepted, the circumstances of this case do not involve the same concerns that the proposition attempts to alleviate.

As the majority’s support indicates, facial challenges carry a heavy burden when the statute may be implemented in a variety of ways. For example, the case cited for support of this proposition, \textit{Washington State Grange v. Washington State Republican Party}, involved a state election regulation that was challenged for its violation of the First Amendment, which, notably, allows the overbreadth doctrine.\textsuperscript{164} Additionally, the case on which \textit{Washington State Grange} rests its support is similarly distinguishable from the challenge here. In \textit{United States v. Salerno},\textsuperscript{165} an individual challenged the constitutionality of his pre-trial detention imposed under the Bail Reform Act. The constitutional challenge here, however, involves a statute’s violation of the Appointments Clause and separation of powers doctrine. Such a challenge presents a situation vastly different from \textit{Washington State Grange} or \textit{Salerno} which involved the implementation of a statute to an individual’s situation.

Although the Fund does present a facial challenge, the majority’s suggestion that “all circumstances” must be considered in determining whether the Act “cannot constitutionally be applied” seems to exaggerate the situation. Here, unlike statutes that may be applied to individuals differently,\textsuperscript{166} the Act either “constrains the President’s ability to see that the laws are faithfully executed” or it does not. The circumstances do not change; they are prescribed by the statute to a specific, known situation. Thus, the constitutionality of this statute turns not on “premature interpretations of [a] statute[ ] in areas where [its] constitutional

\textsuperscript{162} Id. at 684 (citing Wash. State Grange v. Wash. State Republican Party, 128 S. Ct. 1184, 1191 (2008)).


\textsuperscript{164} Wash. State Grange, 128 S. Ct. at 1190 n.6 (“Our cases recognize a second type of facial challenge in the First Amendment context under which a law may be overturned as impermissibly overbroad because a ‘substantial number’ of its applications are unconstitutional, ‘judged in relation to the statute’s plainly legitimate sweep.’” (internal citations omitted)).

\textsuperscript{165} United States v. Salerno, 481 U.S. 739 (1987).

\textsuperscript{166} Id.; see also Wash. State Grange, 128 S. Ct. at 1184.
application might be cloudy;"\textsuperscript{167} rather, it simply turns on the manner in which it is written. Such a challenge does not involve the vast and varied circumstances in which it may be applied, nor does it touch upon the concerns articulated in \textit{Washington State Grange}.\textsuperscript{168}

To the extent that the majority’s opinion was bolstered by this presumption, that weight is unwarranted. Additionally, by ignoring the teaching of Myers and the facts of \textit{Humphrey’s Executor}, the majority incorrectly applied the separation of powers doctrine to the President’s power to remove Board members. Thus, the President’s power to remove the Board members is restricted to the point that it “impede[s] the President’s ability to perform his constitutional duty[.]

\textbf{V. Conclusion}

Corporate officers eagerly awaiting a decision that declares the Board unconstitutional will need to wait longer and possibly forever. Although \textit{Free Enterprise Fund} was incorrectly decided, the opportunity to have the Supreme Court pass on the constitutional questions might be mooted. As Judge Kavanaugh points out, “the constitutional flaws here could be easily and quickly corrected.”\textsuperscript{170} He suggests that Congress could simply amend the statute to require that the President appoint the Board members with the advice and consent of the Senate.\textsuperscript{171} Alternatively, he suggests that Congress simply make the Board part of the SEC “directed, supervised, and removable at will by the Commission.”\textsuperscript{172} Despite these solutions, the Supreme Court has an opportunity to add to its jurisprudence in an area of the law that it has not considered in over twenty years. Even if deciding this case will not significantly alter the regulation of the accounting industry, it could significantly alter the Appointments Clause and the separation of powers doctrine.


\textsuperscript{168} See id. In \textit{Wash. State Grange}, the concerns stated are that facial challenges rest on speculation, are contrary to judicial restraint because they anticipate a constitutional question and formulation an overly broad rule of constitutional law, and they “short circuit the democratic process by preventing laws embodying the will of the people from being implemented in a manner consistent with the Constitution.” \textit{Id.} at 1191.


\textsuperscript{171} \textit{Id.}

\textsuperscript{172} \textit{Id.}
A COSMOPOLITAN APPROACH TO TREATY INTERPRETATION:
WHY SERVICE BY POSTAL CHANNELS SHOULD BE PERMITTED UNDER THE HAGUE CONVENTION

Christine A. Elech∗

I. INTRODUCTION

Approaching the middle of the twentieth century, countries around the world began abandoning their policies of global and judicial isolationism.1 As international commerce, travel, and communication expanded, so too did involvement in international litigation.2 Judicial systems were not prepared for this because there were no internationally acceptable rules for engaging in cross-border civil litigation.3 Since private litigants were responsible for complying with domestic standards as well as the foreign country’s local service laws, it became very difficult to successfully bring an action without a violation of law.4 Many countries began acknowledging the need for an international mechanism to facilitate service of process abroad and eventually met in The Hague to pursue the goal of harmonizing conflict of law principles within private international law. The resulting document is the Convention on the Service Abroad of Judicial and Extrajudicial Documents in Civil or Commercial Matters; its purpose being the simplification and facilitation for the transmission of judicial and extrajudicial documents by private litigants from one state to another.5 The treaty provides private litigants with a number of methods to effect service abroad.6

Currently, the federal circuits are split as to whether service of process abroad on a foreign national by direct mail is proper under the Hague Convention. The specific language in contention is located in Article Ten, stating in part, “[p]rovided the State of destination, does not object, the present

∗ Christine A. Elech is a J.D. candidate for 2010 at Salmon P. Chase College of Law, Northern Kentucky University. She received a B.A. in Diplomacy/Foreign Affairs from Miami University in 2007.

3. Id. at 211.
4. Id.
Convention shall not interfere with . . . the freedom to send judicial documents, by postal channels, directly to persons abroad . . . .”7 On one side of the split, the Fifth and Eighth Circuits conclude that direct mail is not a method of service permitted by the Convention. Under this approach, the Convention authorizes the sending of judicial documents by mail, but only after service of process is accomplished by some other means provided for in the treaty. The reasoning is that the plain meaning of the treaty is clear thus there is no need to analyze legislative intent. On the other side of the split, the Second and Ninth Circuits determine that service by mail is an acceptable means to serve a foreign national. This is founded in the conclusion that, after conducting an analysis of legislative intent, it is clear that the term “send” is intended to include “service.”

This article argues that the Second and Ninth Circuits follow the more persuasive line of reasoning. Part II describes the relevant background on the Hague Convention as well as related information on publications by the Permanent Bureau of the Hague Conference. Part III highlights major cases in all four circuits, demonstrating which side of the split each circuit lies. Part IV analyzes and discusses why the approach taken by the Second and Ninth Circuits is superior and also why such an approach is most beneficial for private litigants. Part V concludes this article.

II. BACKGROUND

The Convention on the Service Abroad of Judicial and Extrajudicial Documents in Civil or Commercial Matters, commonly referred to as the Hague Convention, was adopted in 1965 at the Tenth Session of the Hague Conference on Private International Law.8 The purpose of the Hague Convention is to simplify and facilitate the transmission of judicial and extrajudicial documents by private litigants from one state to another in order to ensure that such documents will be brought to the notice of the addressee in sufficient time.9 The Hague Convention is mandatory where service is made in a signatory nation, and such service must be in compliance with its requirements as well as honor any restrictions a country has imposed on service within its boundaries.10 The Convention provides a plaintiff, authorized to serve process under its country’s laws, with a method to effect service that will give appropriate notice to the foreign defendant.11 As long as the nation concerned has not, in its ratification or in any other part of its law, either explicitly or by compelling implication,
imposed any limits on clauses contained in the Convention, all provisions are applicable.12

The United States Senate ratified the Hague Convention in 1967 and in 1969 it entered into force.13 Because the treaty is Senate-ratified, it is constitutionally declared to be the “supreme Law of the Land,” thus binding all state and federal courts to its decree.14 In the United States, a federal court plaintiff must comply with the provisions of Federal Rule of Civil Procedure Four for serving process abroad, in addition to the Convention’s requirements.15 Rule 4(f)(1) provides that a plaintiff may serve a foreign defendant “by any internationally agreed means of service that is reasonably calculated to give notice, such as those authorized by the Hague Convention . . . .”16 Thus, in Hague Convention signatory countries, service of process abroad must comply with the Convention.

The Convention details multiple channels of transmission in order to service process on a foreign defendant. Articles Two through Six require that member nations designate a Central Authority to receive requests for service, arrange for the transmission of documents to recipients, and return proofs of service.17 Article Five provides in pertinent part that

The Central Authority of the State addressed shall itself serve the document or shall arrange to have it served by an appropriate agency, either- (a) by a method prescribed by its internal law for the service of documents in domestic actions upon persons who are within its territory, or (b) by a particular method requested by the applicant, unless such a method is incompatible with the law of the State addressed.18

While the use of a signatory nation’s Central Authority is commonly called the “main channel of transmission,” it is not the exclusive mode permitted.19 Litigants are allowed to engage in alternative methods as well. For example, Articles Eight and Nine permit the use of diplomatic agents to serve foreign defendants in recipient nations.20 Article Eleven allows two signatories to agree to other methods not specified in the Convention.21 In addition, Article Nineteen allows a country to unilaterally permit other forms of transmission that the Convention does not allow.22 Sections (b) and (c) of Article Ten provide that if

13. See Hague Convention, supra note 5, preamble.
17. See Hague Convention, supra note 5, art. II-VI.
18. Id. at art. V.
19. Permanent Bureau of the Hague Conference on Private Int’l Law, supra note 8, at 35 (noting that there are several alternative channels of transmission).
20. Hague Convention, supra note 5, art. VIII-IX.
21. Id. at art. XI.
22. Id. at art. XIX.
the state of destination does not object, judicial officials may effect service directly through judicial officers in the state of destination, and any person interested in a judicial proceeding may effect service directly through judicial officials in the state of destination. Section (a) of Article Ten stipulates that, provided the state of destination does not object, the Convention will not interfere with the freedom to send judicial documents, by postal channels, directly to persons abroad.

In order to facilitate understanding of the Convention’s basic operation, the Permanent Bureau of the Hague Conference routinely publishes “Practical Handbooks” with information about the Convention and explanation of its language. In reference to Article 10(a), the most recent edition of the handbook states that “if all relevant conditions are fulfilled, transmission of the documents through postal channels includes service of process upon the addressee . . . .” Service by mail is effective if it is allowed by the law of the state of origin and the state of destination has not objected to the use of Article 10(a). A state may notify its objection either when filing its instrument of ratification or accession.

The handbook’s authors also make a point to note the relevant historical context of the language in Article 10(a). The 1965 Convention was the first text drawn up by the Hague Conference dealing with service abroad, as well as the first text to be published in English. The official version is translated into both French and English. The verb “addresser” used in the French version of Article 10(a), translated in English by the verb “send,” has consistently been interpreted as meaning “service” or “notice.” Most importantly, the 2003 Special Commission affirmed the position that the term “send” in the English version of the Convention is to be understood as referring to “service” through postal channels.

The handbook gives detailed insight into specific interpretations intended and preferred by the Convention’s drafters. Also, it provides commentary on various court opinions dealing with ambiguous language. While the published information is not binding law in any American jurisdiction, it is a relevant resource for determining legislative intent.

23. Id. at art. X.
24. Id.
25. PERMANENT BUREAU OF THE HAGUE CONFERENCE ON PRIVATE INT’L LAW, supra note 8, at III.
26. Id. at 69.
27. Id. at 71.
28. Hague Convention, supra note 5, art. XXI.
29. PERMANENT BUREAU OF THE HAGUE CONFERENCE ON PRIVATE INT’L LAW, supra note 8, at 79.
30. See Hague Convention, supra note 5.
31. PERMANENT BUREAU OF THE HAGUE CONFERENCE ON PRIVATE INT’L LAW, supra note 8, at 79.
32. Id. at 151-53.
III. THE SPLIT: U.S. CENTRIC V. COSMOPOLITAN APPROACH

A circuit split exists regarding the interpretation of Article 10(a) of the Hague Convention. This section stipulates that, as long as the state of destination does not object, the Convention will not interfere with the freedom to send judicial documents, by postal channels, directly to persons abroad. Because the Convention enumerates multiple methods of service, courts have reached opposing conclusions as to whether the language is intended to provide an alternative method of service, or if it merely refers to the transmission of judicial documents through direct mail after service has been accomplished by other means. The United States Court of Appeals for the Fifth and Eighth Circuits take strict approaches and base their conclusions on traditional rules of statutory construction. These courts hold that the word “send” is not the equivalent of service of process because the word “service” is used in other sections of the Convention, all referring to the forwarding of documents for the purpose of service. The United States Court of Appeals for the Ninth and Second Circuits, on the other hand, look to the broad purpose and history of the Hague Convention and conclude that, because the Convention’s purpose is to facilitate and simplify service abroad and also because there are sources stipulating the authors’ intent, the term “send” is intended to mean “send documents for the purpose of service.”

A. The U.S. Centric Approach

Two circuits pioneer the approach centered in traditional statutory construction. The most noted case advocating such an approach is Bankston v. Toyota, decided by the Eighth Circuit in 1989. In 2002, the Fifth Circuit subscribed to Bankston’s rationale with Nuovo Pignone, SpA v. Storman.

In Bankston, the plaintiffs filed suit in the United States District Court for the Western District of Arkansas against Toyota Motor Corporation seeking damages resulting from an accident involving a Toyota truck. Toyota is a Japanese automobile manufacturer. The plaintiffs made two attempts to serve process, the first made on an affiliated United States corporation in California, as Toyota’s claimed agent. This attempt was deemed improper and the plaintiffs

34. Hague Convention, supra note 5, art. X.
35. See Nuovo Pignone, SpA v. Storman Asia M/V, 310 F.3d 374, 384 (5th Cir. 2002); Bankston v. Toyota Motor Corp., 889 F.2d 172, 174 (8th Cir. 1989).
36. Nuovo Pignone, SpA, 310 F.3d at 383-84; Bankston, 889 F.2d at 173-74.
37. Brockmeyer v. May, 383 F.3d 798, 802-03 (9th Cir. 2004); Ackermann v. Levine, 788 F.2d 830, 839-840 (2d Cir. 1986).
38. Bankston, 889 F.2d at 172.
39. Id.
40. Id.
were given time to serve Toyota in accordance with the Hague Convention.\textsuperscript{41} The next attempt to serve process was made through registered mail to Tokyo, Japan, prompting Toyota to move to dismiss based on the argument that such a method of service still did not comply with the Convention.\textsuperscript{42} The District Court ruled for Toyota by concluding that Article 10(a) does not permit service of process by registered mail.\textsuperscript{43} On appeal, The United States Court of Appeals for the Eighth Circuit affirmed the District Court.\textsuperscript{44}

First, the appellate court observed that the word "service" is specifically used in other sections throughout the Convention, including 10(b) and 10(c).\textsuperscript{45} If the authors of the Convention intended for Article 10(a) to allow an alternative manner to serve process, the word "service" would be substituted for "send."\textsuperscript{46} The court reasoned that the provision must only provide a method of sending documents after service of process is accomplished by other means.\textsuperscript{47}

The primary basis for the court’s conclusion lies in traditional statutory construction. Quoting United States Supreme Court cases, the Eighth Circuit stipulated that "[i]t is a ‘familiar canon of statutory construction that the starting point for interpreting a statute is the language of the statute itself. Absent a clearly expressed legislative intention to the contrary, that language must ordinarily be regarded as conclusive.’\textsuperscript{48} Also, the court noted that "where a legislative body ‘includes particular language in one section of a statute but omits it in another section of the same Act, it is generally presumed that [the legislative body] acts intentionally and purposely in the disparate inclusion or exclusion.’\textsuperscript{49}

In 2002, the Fifth Circuit joined the Eighth in \textit{Nuovo Pignone, SpA v. Storman}. Nuovo Pignone, SpA, an Italian company, contracted with Fagioli, a company providing global transportation for heavy lift cargo, in order to transport a reactor from Italy to Louisiana.\textsuperscript{50} The reactor was damaged while being transferred to a barge in New Orleans and Nuovo brought suit in tort and contract.\textsuperscript{51} Nuovo effected service of process by sending the complaint and summons by Federal Express mail to Milan, Italy.\textsuperscript{52} The United States District Court for the Eastern District of Louisiana concluded that service by mail of

\begin{itemize}
\item \textsuperscript{41} Id.
\item \textsuperscript{42} Id.
\item \textsuperscript{43} Id.
\item \textsuperscript{44} Bankston, 889 F.2d at 174.
\item \textsuperscript{45} Id. at 173.
\item \textsuperscript{46} Id. at 173-74.
\item \textsuperscript{47} Id. at 174.
\item \textsuperscript{48} Id. (quoting Consumer Prod. Safety Comm’n v. GTE Sylvania, Inc., 447 U.S. 102, 108 (1980)).
\item \textsuperscript{49} Id. (quoting Russello v. United States, 464 U.S. 16, 23 (1983)).
\item \textsuperscript{50} \textit{Nuovo Pignone, SpA}, 310 F.3d at 377.
\item \textsuperscript{51} Id.
\item \textsuperscript{52} Id. at 377-78.
\end{itemize}
foreign parties is permissible under Article 10(a) of the Hague Convention.53 On appeal, the United States Court of Appeals for the Fifth Circuit reversed the District Court’s judgment.54

At the appellate level, Fagioli argued that service by mail violates Rule 4(f)(1) of the Federal Rules of Civil Procedure, which permits service of process on a foreign corporation “by any internationally agreed means reasonably calculated to give notice, such as those means authorized by the Hague Convention” and service by mail is not included.55 First, the court noted that nowhere else in the Hague Convention is the word “send” used to refer to service of process; instead the authors use the words “serve,” “service,” and “to effect service” in other sections.56 Such sections referring to the forwarding of documents for the purpose of service include (b) and (c) of Article Ten, and Articles Nine, Fifteen, and Sixteen.57 The court determined that if the drafters of the Convention wished to allow service by mail, Article 10(a) would read “service” and not “send.”58 Also, the court mentioned that other provisions in the Convention describe more reliable methods of effecting service than regular mail.59

Second, the court in Nuovo Pignone, SpA echoes Bankston by relying on principles of statutory interpretation.60 The court declared that without a clearly expressed legislative intention to the contrary, a statute’s language must be regarded as conclusive.61 In addition, the court noted that where a legislative body includes particular language in one section of an act but omits it in another section, it is presumed that it was intentional in the disparate inclusion or exclusion.62

Third, as further evidence that service by mail is not allowed under the Hague Convention, the court pointed to the Advisory Committee notes of the 1963 amendments to Rule Four of the Federal Rules of Civil Procedure.63 The Court emphasized that these notes, which recognize that service of process beyond the territorial limits of the United States may involve difficulties not encountered in domestic service, evidence the fact that service by mail does not further the Convention’s goal of ensuring the delivery of service.64

Fourth, Nuovo points out that the Hague Convention’s purpose is not only to simplify service of process, but to ensure that notice is delivered to foreign

53. Id. at 378.
54. Id. at 385.
55. Id. at 383 (quoting FED. R. CIV. P. 4(f)(1)).
57. Id. at 384.
58. Id.
59. Id. at 384-85.
60. Id. at 384.
61. Id.
63. Id. at 384 n.17.
64. Id. at 384.
addressees in sufficient time. The court was not confident that regular mail service is sufficient enough to guarantee receipt in sufficient time and stated that it is unlikely that the drafters would have put in place so many other reliable methods of effective service, while simultaneously permitting the uncertainties of service by mail. For all of the above reasons, the court determined that the drafters purposefully elected to use forms of the word “service” throughout the Convention, and intentionally confined use of the word “send” to Article 10(a).

The U.S. Centric approach taken by the Eighth and Fifth Circuits is founded in the use of general principles of statutory construction. The approach gives great deference to the final product of the Convention, intending to support the explicit language ratified by the Senate.

B. The Cosmopolitan Approach

Two circuits spearhead the approach based primarily in the historical context of the Hague Convention and international sources of interpretation. The most prominent case advocating such an approach is *Ackermann v. Levine*, decided by the Second Circuit in 1986. Since then, the Ninth Circuit has followed *Ackermann’s* lead with *Brockmeyer v. May* in 2004.

In *Ackermann v. Levine*, a German plaintiff filed suit in Germany and served process on an American defendant by registered mail. The plaintiff, Peter Ackermann, was a German citizen practicing law in West Berlin. The defendant, Ira Levine, was an American citizen engaged in the real estate business in New York. Levine met a West German business promoter, “Bauer,” and subsequently became interested in recruiting German investors in a proposed real estate project. Once in Germany, Bauer recommended that Levine talk with his friend and attorney, Peter Ackermann. Levine and Ackermann met to discuss legal and tax issues, but the parties never addressed the nature of their relationship. Ackermann then presented Levine’s project to German financial institutions, but his efforts failed. Levine then went to Germany and negotiated a deal without Ackermann. Ackermann sent a bill to Levine for the legal services he provided, which Levine never paid.

65. Id.
66. Id.
67. Id. at 385.
68. Ackermann v. Levine, 788 F.2d 830, 834, 837 (2d Cir. 1986).
69. Id. at 834.
70. Id.
71. Id.
72. Id.
73. Id. at 835.
74. *Ackermann*, 788 F.2d at 836.
75. Id.
76. Id.
Suit was initiated in the Regional Court of Berlin to recover legal fees. The German court sent a summons and complaint to the German Consulate in New York, which mailed them by registered mail to Levine’s former address. Levine did not answer, and the court entered a default judgment against him.

Ackermann then commenced an action seeking enforcement of the foreign judgment in the United States District Court for the Southern District of New York. The court declined to enforce the judgment in part because service of process by registered mail violated the Hague Convention. The United States Court of Appeals for the Second Circuit disagreed with the district court and approved Ackermann’s service of process by registered mail. Also, the court upheld the foreign judgment. The court relied primarily on the purpose and history of the Convention to interpret the word “send” in Article 10(a) as including the meaning “serve.”

The court determined that Ackermann used Article Eight to effect service directly through diplomatic or consular agents. Ackermann sent the summons and complaint to the German Consulate in New York who then forwarded the documents by registered mail to Levine. In its analysis, the Second Circuit first stated that the word “send” in Article 10(a) is intended to mean “service.” The Convention’s “reference to ‘the freedom to send judicial documents by postal channels, directly to persons abroad’ would be superfluous unless it was related to the sending of such documents for the purpose of service.” The use of the term “send” instead of “serve” was determined to be careless drafting. The court also noted that Article Ten of the Convention states that sending judicial documents by postal channels is legitimate as long as the state of destination does not object. Because the United States has made no objection to the use of “postal channels” under Article 10(a), the court deemed service to be effective.

**Brockmeyer v. May**, decided by the Ninth Circuit, is another case supporting the broad construction of the Convention followed by the Second Circuit. In

---

77. Id. at 837.
78. Id.
79. Id.
80. Ackermann, 788 F.2d at 837.
81. Id. at 845.
82. Id. (upholding foreign judgment, but declining to enforce compensation for the “study of project files” because “there was no evidence of authorization or work product”).
83. Id. at 838-40.
84. Id. at 839.
85. Id.
86. Ackermann, 788 F.2d at 839.
87. Id. (quoting Shoei Kako v. Superior Court, 33 Cal. App. 3d 808, 821-22 (1973)).
88. Id.
89. Id.
90. Id.
Brockmeyer, an American plaintiff served process by ordinary international mail on a British corporation at its London post office box address.91

The dispute in Brockmeyer arose out of a trademark infringement lawsuit filed by Plaintiff Ronald Brockmeyer against several defendants including Marquis Publications, Ltd., a company registered under British law.92

Brockmeyer sued Marquis in the United States District Court for the Central District of California to enjoin its use of his trademarked symbol.93 Plaintiff made two attempts to serve Marquis.94 On the first attempt, Brockmeyer sent a copy of the summons and complaint, along with a request for waiver of service, by ordinary first class mail.95 The second attempt included sending the summons and complaint by first class mail to the same post office box.96 The district court entered a default judgment against Marquis and denied its motion to set it aside.97

The determination of the United States Court of Appeals for the Ninth Circuit depended on whether Marquis was properly served under the Hague Convention.98 Since Brockmeyer used first class mail as his attempted method for service, Article 10(a) was applicable. The court joined the Second Circuit in holding that the meaning of “send” in Article 10(a) includes “serve.”99

The court first referred to court opinions of Hague Convention member states including Canada and the European Court of Justice concluding that the meaning of “send” includes “serve.”100 Next the court relied heavily on Ackermann’s reasoning founded in the Convention’s purpose to facilitate international service of judicial documents. The court also looked to commentaries on the history of negotiations.101 According to the official report of the negotiations, the first paragraph of Article Ten of the draft Convention, which “except for minor editorial changes” is identical to Article Ten of the final Convention, was intended to permit service by mail.102

Also, an earlier published handbook states that “to interpret Article 10(a) not to permit service by mail would ‘contradict what seems to have been the implicit understanding of the delegates at the 1977 Special Commission meeting, and indeed of the legal

91. Brockmeyer v. May, 383 F.3d 798, 799 (9th Cir. 2004).
92. Id. at 800.
93. Id.
94. Id.
95. Id.
96. Id.
97. Brockmeyer, 383 F.3d at 800-01.
98. Id. at 801.
99. Id. at 802.
101. Id. at 803-04.
102. Id. at 802-03 (citing 1 BRUNO A. RISTAU, INTERNATIONAL JUDICIAL ASSISTANCE (CIVIL AND COMMERCIAL) § 4-3-5, at 204-05 (2000)).
literature on the Convention and its predecessor treaties.” Additionally, the United States delegate to the Hague Convention reported to Congress that Article 10(a) allowed service by mail.

As further evidence, the court considered the specific disapproval of the Eighth Circuit’s decision in the Bankston case by the United States Department of State. This disapproval is evidenced by a letter written by the Deputy Legal Advisor of the State Department to the Administrative Office of United States Courts. Finally, the court looked to State Department circulars which indicate that service by mail is permitted in international civil litigation.

The Second and Ninth Circuits base their interpretations of Article 10(a) on a significant amount of persuasive authority rather than focus on general principles of statutory construction.

IV. ANALYSIS

An interpretation of Article 10(a) that permits service of process by direct mail is the preferable approach to adopt for the following reasons. First, the authors of the Convention have published that their intentions at execution, in conformance with the explicit purpose of the Convention, were to permit service of process under Article 10(a). Second, there are a number of opinions issued by signatory nations which are in line with the authors’ intent as well as the approach taken by the Ninth and Second Circuits. Third, the United States Executive Branch understands the provision to allow service of process by postal channels. Fourth, because there are established rules of treaty interpretation, courts should give deference to them rather than limit an analysis to general principles of statutory construction. Finally, service by postal channels provides private litigants significant advantages including low costs and relatively fast transmission. Thus, adopting a cosmopolitan approach to the interpretation of international treaties is likely to produce conclusions coinciding with legislative intent as well as result in international consistency.

A. Purpose and Legislative Intent

The Preamble of the Hague Convention states that a chief purpose of its existence is to “improve the organisation [sic] of mutual judicial assistance…by

104. Id. (citing S. Exec. R. No. 6-90, at 13 (1967) (statement by Philip W. Amram, U.S. delegate to the Hague Convention)).
105. Id.
106. Id.
simplifying and expediting the procedure . . . ”.108 This purpose is reflected by the numerous methods of service available to private litigants. While the Convention clearly details those methods, there are external sources directly affiliated with the Permanent Bureau of the Hague Conference that can aid a court when addressing a question of interpretation.

In order to facilitate an accurate understanding of the Convention’s basic operation, the Permanent Bureau of the Hague Conference has routinely published “Practical Handbooks” with information about the Convention and explanation of its language.109 The most recent edition of the handbook, published in 2006, specifically addresses the circuit split at issue in this article and endorses the approach taken by the Ninth and Second Circuits.

First, the handbook is completely unambiguous when explaining the meaning of Article Ten. It states that “if all the relevant conditions are fulfilled, transmission of the documents through postal channels includes service of process upon the addressee . . . .”110 The relevant conditions include (a) whether service by mail is allowed by the law of the state of origin and all the conditions imposed by that law for service by mail have been met, and (b) the state of destination has not objected to the use of Article 10(a).111 All conditions are fulfilled in the United States whether the country acts as a state of origin or as a state of destination. As a state of origin, the Federal Rules of Civil Procedure provide that a private litigant must accomplish service of process abroad according to the Convention.112 Thus, as long as the Convention is interpreted to allow service by mail, this condition is fulfilled when serving process in this way. As the state of destination, the United States must allow service on a private litigant since the Senate did not make any specific objections to the provisions of the Convention at the time of ratification.113 Thus, this condition is also fulfilled.

Second, the handbook notes the importance of the Convention’s historical context. The Hague Convention was the first text drawn up by the Hague Conference that (1) dealt with service abroad and (2) was translated into an official English version.114 In the official French version of this Convention, the verb “addresser” is used in Article 10(a).115 In French, “addresser” is not the equivalent of “service” but, in accordance with the authors’ intent, it has

108. Hague Convention, supra note 5, preamble.
109. PERMANENT BUREAU OF THE HAGUE CONFERENCE ON PRIVATE INT’L LAW, supra note 8, at III.
110. Id. at 69.
111. Id. at 71.
113. See Hague Convention, supra note 5, art. XXI.
114. PERMANENT BUREAU OF THE HAGUE CONFERENCE ON PRIVATE INT’L LAW, supra note 8, at 79.
115. Id.
consistently been interpreted as meaning “service” or “notice.” The verb “addresser” was used in the same context in previous Conventions concerning civil procedure, all of which were written solely in French. The English version of Article 10(a) uses the word “send.” Just as “addresser,” “send” is not the equivalent of “service.” It is a rule of treaty interpretation that official versions of a treaty in different languages are assumed to have the same meaning, and effort should be made to construe each version accordingly. Because “addresser” has been construed to mean “service” in three predecessor treaties concerning civil procedure, and “send” is a translation from “addresser” for purposes of the first English version, “send” should be construed the same way “addresser” has been.

Third, the intent of the drafters is also clear in the Convention’s official report. It stipulates that Article 10(a) of the Convention’s final draft is identical to the initial draft, except for minor editorial changes. The report notes that the initial draft permits service by mail. If only minor editorial changes were made between the initial and final drafts, it is not difficult to conceive that the word “send” in the final draft had the identical meaning as it had in the initial draft. This report is persuasive when taken into consideration with the number of other concurring sources.

Fourth, the Special Commission explicitly confirmed that Article 10(a) permits service by mail. In 2003, the Commission met in The Hague to review the practical operation of multiple Hague Conventions. This meeting was attended by 116 delegates representing 57 member states, including the United States. At the meeting’s completion, the Permanent Bureau of the Hague Conference on Private International Law published the Commission’s “Conclusions and Recommendations.” This information has been made public in order to be a global resource for courts when tackling interpretation issues originating from the Convention’s language, just as the Practical Handbook is also intended to be a resource. Regarding the Convention on service, the

116. Id.
117. Id.
118. Hague Convention, supra note 5, art. X.
120. Brockmeyer, 383 F.3d at 802-03 (citing 1 BRUNO A. RISTAU, INTERNATIONAL JUDICIAL ASSISTANCE (CIVIL AND COMMERCIAL) § 4-3-5, at 204-05 (2000)).
121. Id. (citing 1 BRUNO A. RISTAU, INTERNATIONAL JUDICIAL ASSISTANCE (CIVIL AND COMMERCIAL) § 4-3-5, at 204-05 (2000)).
123. Id.
124. Id. at n.1.
125. Id.
Commission affirmed its “clear understanding” that the term “send” in Article 10(a) is to be understood as meaning “service.” Because the United States attended the 2003 Special Commission meeting at The Hague, the American delegates were fully aware of the Commission’s results. If the delegates did not support certain portions of the literature, it would be noted within the text. However, no such footnotes exist.

While the sources described above are not binding law in any American jurisdiction, they are legitimate secondary sources available for interpretation of the Convention’s language and ascertaining legislative intent. In very few situations are there publications dedicated to aiding both litigants and courts in the interpretation of statutory language. Complete disregard of these resources implies a disinterest in issuing a judgment that accurately reflects legislative intent.

B. Foreign Adjudication

A variety of foreign signatory nations have echoed the interpretation of “send” as including “service.” In *Brockmeyer*, the court gave significant deference to foreign adjudication on the interpretation of Article 10(a). The two most relevant opinions cited include one issued by the European Union Court of Justice and one by a Canadian court.

The European Union Court of Justice stipulates that Article 10(a) allows service by post. This is important because the European Court of Justice declares European Community law to be supreme over conflicting national law. In essence, the highest European court has ruled that all twenty seven member states should interpret Article 10(a) of the Hague Convention to allow service by post, regardless of any conflicting national law. Therefore, as long as a member-state fulfills the requisite conditions detailed by the Practical Handbook, it should allow service abroad by postal channels. This decision is significant because many European Union member states are also Hague Convention signatories. If all relevant European states recognize the European

---

126. Id. at 11.
127. See id. at 10 n.3-5 (noting that when Russia did not support a position, a footnote was inserted to reflect its opposition).
129. Case C-412/97, E.D. Srl. v. Italo Fenocchio, 1999 E.C.R. I-3845 (“The national court adds that that is particularly true in the case of States which are signatories to the Hague Convention of 15 November 1965 on the Service Abroad of Judicial and Extrajudicial Documents in Civil or Commercial Matters, Article 10(a) of which allows service by post.”).
131. Hague Convention, *supra* note 5, art. II-VI.
Court of Justice’s decision, the Fifth and Eighth Circuits are left defying what is virtually an international consensus on the interpretation of Article 10(a).

In Canada, the Judicial District of Calgary noted that “[t]he starting point of statutory construction is the words of the statute. If they are clear, that is the end of the matter.”133 If however, a court is confronted with a phrase that it can interpret in more than one way, it must consider the words of the provision “in their entire context and in their grammatical and ordinary sense harmoniously with the scheme of the Act; the object of the Act, and the intention of Parliament.”134 If the Canadian court believed that the word “send” is clear, it would have ruled that service was not included. Because it did not rule this way, an analysis into legislative intent was likely conducted.

Although the Canadian judicial system supports the same principles of statutory construction as the United States, this court did not end their analysis with an English understanding of the word “send,” as did the Fifth and Eighth Circuits. Such a deviation from tradition likely resulted because the Hague Convention is not written in English by a Canadian legislature. Because international treaties are often filled with latent ambiguities, the Canadian court placed great weight on legislative intent in order to construe the Hague Convention as accurately as possible.

American courts that are refuting the dominant view of this issue must look beyond the traditional limits on statutory construction, as has Canada, and recognize that abandonment of state-centric rules will likely produce a precise result.

C. The Executive Branch

The Executive Branch of the United States Government has made it clear that a correct interpretation of Article 10(a) of the Hague Convention includes service by postal channels. Its understanding of the Hague Convention is important because the Department of State plays a direct role in negotiations of international agreements. No other branch is approached in the treaty making process until a treaty is considered for Senate ratification.135 An example of

---

134. Id. at ¶ 59.
deep Executive Branch involvement in this Convention is its power to designate the Central Authority detailed in Articles Two through Seven.136  The Central Authority has always been an Executive Branch function, with both the Department of State137 and the Department of Justice employing the role.138 Thus, any private litigant choosing to effect service abroad through the Central Authority must comply with Executive Branch procedural requirements. This branch was not only involved in the development of the treaty but also plays an important role in its implementation.

While United States courts have final authority to interpret international treaties for purposes of their application as law, the Supreme Court has held that it should give weight to interpretations made by the Executive Branch.139 There is a significant amount of undisputable evidence demonstrating the Executive Branch’s understanding of Article 10(a). First, Philip Amram, the United States delegate to the Hague Convention, reported to Congress that Article 10(a) permits service by mail.140 This is significant not only because Mr. Amram was appointed by the Executive Branch, but also because the information was relayed by him to Congress in 1967, the year of ratification. Therefore, at the time of voting, all Senators were aware of the Executive Branch’s understanding resulting from Convention negotiations. Second, State Department Circulars indicate that service by mail is permitted in international civil litigation.141 Regarding service by postal channels, the circular on the Hague Convention stipulates that international mail service of civil summons is proper unless a State has entered an appropriate reservation under Article Ten.142 Because this information is derived from a 2003 Circular, it reiterates that the Executive Branch’s understanding of Article 10(a) has remained constant since 1967.

Also, the State Department has disapproved of the decision in Bankston via a letter from Alan J. Kreczko, United States Department of State Legal Advisor, to

136. See supra Part II.
137. Hague Convention, supra note 5, at n.1 (“[T]he United States Department of State is designated as the Central Authority . . . .”).
139. El Al Isr. Airlines, Ltd. v. Tsui Yuan Tseng, 525 U.S. 155, 168 (1999) (“Respect is ordinarily due the reasonable views of the Executive Branch concerning the meaning of an international treaty.”). See also Kolovrat v. Oregon, 366 U.S. 187, 194 (1961) (“While courts interpret treaties for themselves, the meaning given them by the departments of government particularly charged with their negotiation and enforcement is given great weight.”); Factor v. Laubenheimer, 290 U.S. 276, 295 (1933) (“[I]n resolving doubts the construction of a treaty by the political department of the government, while not conclusive upon courts called upon to construe it, is nevertheless of weight.”); RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW OF THE UNITED STATES § 326(2) (1987) (“Courts in the United States have final authority to interpret an international agreement for purposes of applying it as law in the United States, but will give great weight to an interpretation made by the Executive Branch.”).
140. Brockmeyer, 383 F.3d at 803 (citing S. Exec. R. No. 6-90, at 13 (1967) (statement by Philip W. Amram, U.S. delegate to the Hague Convention)).
141. Id. (citing U.S. Dep’t of State, Circular, supra note 107).
142. U.S. Dep’t of State, Circular, supra note 107.
the Administrative Office of United States Courts and the National Center for State Courts. Mr. Kreczko stipulates that once the Eighth Circuit decision was brought to the attention of the Office of the Legal Advisor in the Department of State and the Office of Foreign Litigation in the Justice Department, which at the time served as the Central Authority, both departments immediately took issue with it. The Office of the Legal Advisor understands the facts in Bankston to suggest “a judgment by a court in the United States based on service on the defendant in Japan by mail, while capable of recognition and enforcement throughout the United States, may well not be capable of recognition and enforcement in Japan by the courts of that country.” The Japanese statement suggests that service by mail is not considered valid under Japanese law. Thus, if Japan, as the state of destination, did allow such service, an American plaintiff would be permitted to effect service by mail on a Japanese defendant because it is not prohibited by American law. Therefore, the decision in Bankston is incorrect to the extent that it suggests that the Hague Convention does not permit, as a method of service, the sending of a copy of a summons and complaint by mail to a foreign defendant.

The Executive Branch should share the authority to interpret international agreements since the President is the country’s “sole organ” in its international relations and is responsible for carrying out agreements with other nations. The execution of an international agreement as domestic law often requires interpretation of the agreement, and the President determines what the agreement means in the first instance. Because the Executive Branch’s role is fundamental to an agreement’s execution, courts should give deference to its understanding and interpretation. In this case, the Fifth and Eighth Circuits are defying the Supreme Court by not considering the Executive branches unambiguous understanding of Article 10(a).

D. Principles of Statutory Construction and International Treaties

The arguments made by the Fifth and Eighth Circuits rely upon traditional canons of statutory construction. Although such principles often produce

143. See The Am. Soc’y of Int’l Law, United States Department of State Opinion Regarding the Bankston Case and Service by Mail to Japan Under the Hague Convention, 30 Int’l Legal Materials 260, 260 (1991) (“The United States Department of State has recently stated that it believes the decision to be incorrect.”).
144. Id. at 260-61 (quoting Letter from Alan J. Kreczko to the Admin. Office of U.S. Cts. (Mar. 14, 1991)).
145. Id. at 261 (quoting Letter from Alan J. Kreczko to the Admin. Office of U.S. Cts. (Mar. 14, 1991)).
146. Id. (quoting Letter from Alan J. Kreczko to the Admin. Office of U.S. Cts. (Mar. 14, 1991)).
147. Brockmeyer, 383 F.3d at 803.
149. Id. at cmt. b.
accurate interpretations of American legislation, courts should not exclusively use them when interpreting international treaties. Such principles may be starting points for statutory construction, but courts must go further. In fact, there are principles advocated by the Supreme Court which focus solely on treaties. If such rules had been considered by the courts in Bankston and Nuovo Pignone, SpA, the decisions reached would most likely be different.

The first canon relied upon by Bankston and Nuovo Pignone, SpA, as stated by the Supreme Court, is that “[t]he starting point for interpreting a statute is the language of the statute itself. Absent a clearly expressed legislative intention to the contrary, that language must ordinarily be regarded as conclusive.”150 This principle is not sufficient to thoroughly interpret Article 10(a) because it does not require a deep inquiry into legislative intent. If the rule provided a lower standard to show intent contrary to the treaty’s language, these courts would have looked further than the provision itself. A further inquiry would include a number of available sources dating back to 1967, each which detail a consistent meaning of Article 10(a). These sources include, but are not limited to, Philip Amram’s report to Congress, the Convention’s official report, the 2003 Special Commission’s Conclusions and Recommendations, and the Practical Handbook.151

The other general principle of statutory construction relied upon by the Fifth and Eighth Circuits refers to the instance where a legislative body includes particular language in one section of a statute but omits it in another section of the same act.152 In such a case, it is generally presumed that the legislative body acts intentionally and purposefully in the disparate inclusion or exclusion.153 While this principle seems to clarify Article 10(a), because “service” is used in other provisions, it should not be used as a rule of thumb for treaties written in a number of languages.

The Hague Convention has official versions written in both English and French,154 and neither is supreme to the other. Because the two versions cannot possibly be identical, a court should consider both. In the English version, the term “service” is clearly omitted from Article 10(a) and yet included in other provisions providing methods of service.155 In the French version, the verb “adresser” is used in Article 10(a).156 “Adresser” is the French equivalent of the

151. See supra Part IV.
154. See Hague Convention, supra note 8, at 79.
English verb “to send.” Thus, for purposes of translation, “adresser” and “send” correctly mirror one another. In French, however, “adresser” has consistently been interpreted as meaning “service” or “notice.” Thus, “send” and “service” are legal synonyms in French. The drafters, who had not written an official English version of any Hague Convention until this point, may have determined that because “adresser” and “send” are translational equivalents, and “adresser” is interpreted as “service” in French, using “send” in the English version would also be interpreted to include “service.” Thus, simply because the English version contains “service” in provisions other than Article 10(a), no presumption that the omission was intentional should be made. Such a general principle of statutory construction should not apply to treaties officially documented in two or more languages.

According to the Supreme Court, the primary rule of construction is to ascertain and declare the intention of the legislature. A construction should not be adopted to nullify, destroy, or defeat the intention of the legislature, and general rules of statutory construction cannot always produce accurate interpretations. The ability to rely on additional principles when confronted with treaty interpretation enables a court to engage in a deeper analysis. One principle is that treaties are to be interpreted in accordance with the ordinary meaning given to its terms in their context and in the light of its object and purpose. This principle indicates that courts must look beyond the language of the treaty itself. With regards to the Hague Convention, its object and purpose are explicitly detailed as “simplifying and expediting” service abroad. The ability to use postal channels undoubtedly simplifies and expedites service of process because documents are sent directly to the address listed. A private litigant does not have to hassle with a Central Authority or diplomatic agents, which work more slowly because documents must be processed through bureaucratic entities prior to transmission.

An additional principle is that the plain meaning of a treaty applies unless its application would be inconsistent with the intent and expectations of the signatories. “Unless” is the key word in this rule. In the case of Article 10(a), the Executive Branch, as well as courts of signatory nations have expressed that they understand the provision to include service of process. The Executive

157. Id.
158. Id.
160. Id.
161. 2 NANDA & PANSIUS, supra note 119, § 10:5 (citing RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW OF THE UNITED STATES § 325(1) (1987)).
162. Hague Convention, supra note 5, preamble.
163. 2 NANDA & PANSIUS, supra note 119, § 10:5 (citing U.S. v. Stuart, 489 U.S. 353, 365 (1989)).
164. See supra Part IV.
Branch has the authority to negotiate and sign international agreements; thus, the Judicial Branch should give deference to the Executive’s expectations. Additionally, in many other member-states it is understood that Article 10(a) permits service of process by postal channels. For these reasons, the plain meaning of this provision cannot apply.

The most persuasive principle of treaty interpretation is that treaties should be liberally construed consistent with their intent; thus, courts may look beyond the text of a treaty to its history, its negotiations, and the practical construction adopted by the parties. This principle directly reflects this article’s premise in that courts must construe treaties from a cosmopolitan-like standpoint. First, the rule reiterates a former principle by stating that courts should consider the interpretation of other parties. Clearly, other member-states have construed Article 10(a) to permit service abroad by postal channels. Additionally, this rule indicates that courts should look to a treaty’s history and negotiations. The Convention’s history is evidenced by its purpose to simplify and expedite service of process abroad, and its negotiations reflect the delegates’ intent to permit service by postal channels. First, the Convention’s official report stipulates that the Article 10(a) of the Convention’s final draft is virtually identical to the initial draft which intended to permit service by mail. Additionally, Philip Amram, the American delegate to the Convention, who had direct involvement in negotiations, reported identical information to Congress.

E. Advantages of a Cosmopolitan Approach

An interpretation of Article 10(a) that includes service of process offers an important advantage from the standpoint of a litigant attempting to effect service. The ability to serve process by regular mail aids in judicial economy. The Federal Rules of Civil Procedure are intended to be “construed and administered to secure the just, speedy, and inexpensive determination of every action and

166. See supra Part IV.
168. See supra Part IV.
169. Hague Convention, supra note 5, preamble.
170. Brockmeyer, 383 F.3d at 802-03 (citing BRUNO A. RISTAU, INTERNATIONAL JUDICIAL ASSISTANCE (CIVIL AND COMMERCIAL) §§ 4-3-5, at 204-05 (2000)).
171. Id. at 803 (citing S. Exec. R. No. 6-90, at 13 (1967) (statement by Philip W. Amram, U.S. delegate to the Hague Convention)).
proceeding.” This is relevant to the Hague Convention because service of process by postal channels promotes the goal by saving both time and money.

Although the Convention allows multiple methods of service, the use of postal channels is faster than use of a country’s Central Authority. The Convention does not establish a specific time period within which a Central Authority must effect service, although service within “a reasonable time period” is assumed. Unfortunately though, the speed of service by Central Authorities varies significantly by country. In the United States, service is often effected within weeks, but in Europe it often takes months, and in Asia and Latin America service may not ever be effected. This delay is most likely due to the bureaucratic nature of government agencies and different cultural concepts regarding the importance of time. Direct service by postal channels can be effected within a matter of days. Also, if the term “postal channels” includes modern methods, such as fax and email, service may be effected instantly.

Service by postal channels can also save a private litigant a significant amount of money. According to Article Five of the Convention, “If the document is to be served . . . the Central Authority may require the document to be written in, or translated into, the official language or one of the official languages of the State addressed.” This requirement includes translation fees. Also, Article Twelve states that an applicant must pay costs resulting from the employment of a judicial officer and the use of a particular method of service. If a private litigant serves process through direct mail, none of the above fees are applicable. If he serves process by regular mail, only postal charges will apply, and if by fax or email, it will likely be free. Thus, from the standpoint of a private litigant, the ability to serve process by mail offers great advantages. A plaintiff can avoid accumulating more costs than necessary, prevent suffering additional damages, and may have the ability to effect service before a relevant statute of limitations tolls.

Taking a cosmopolitan approach to interpretation of the Hague Convention will enable courts to conclude that Article 10(a) permits service of process by direct mail. This conclusion coincides with the authors’ intention to permit service abroad by postal channels as well as a number of court opinions issued by signatory nations reaching the same result. Additionally, the United States Executive Branch has the authority to negotiation and execute international

173. See supra Part II.
175. Id. at 219.
176. Id.
177. Born, supra note 1, at 811.
178. Hague Convention, supra note 5, art. V.
179. Hague Convention, supra note 5, art. XII.
180. Hawkins, supra note 2, at 218.
treaties, and it understands the Article 10(a) to permit service. Further, if a court considers established principles of treaty interpretation, its analysis of legislative intent is likely to reach an accurate conclusion. Finally, the ability for a private litigant to effect service abroad presents significant advantages including low costs and relatively fast transmission. Such advantages promote judicial efficiency and economy.

V. CONCLUSION

A circuit split exists regarding whether service of process abroad on a foreign national by direct mail is proper under the Hague Convention. Courts on one side of the split adopt a U.S. centric approach refusing to permit service by direct mail because, according to general principles of statutory construction, the term “send” is clear on its face. However, courts on the other side of the split take a cosmopolitan approach to the interpretation of Article 10(a) by conducting an analysis of legislative intent resulting in a determination that “send” is actually intended to include “service.”

Despite being a relatively laborious process, the cosmopolitan approach taken by the Second and Ninth Circuits should be adopted across the circuits for the following reasons. First, it is clear that the intention of the authors is to permit service of process under Article 10(a). Also, there are judicial opinions issued by other member-states concluding that service by mail is authorized. Further, the United States Executive Branch understands Article 10(a) to allow service abroad by postal channels.

In addition to a large consensus on the interpretation of Article 10(a), courts should be further persuaded by existing rules of treaty interpretation. Because such principles have been formulated specifically for the interpretation of treaties, courts should give them deference over general principles of statutory construction. Moreover, courts should permit service by postal channels for practical reasons because private litigants are provided with significant advantages such as low cost and the quick transmission of documents. Therefore, a cosmopolitan approach to treaty interpretation provides courts with a better opportunity to construe the Hague Convention in accordance with its legislative intent.
I. INTRODUCTION

In May of 2008, the United States Court of Appeals for the Second Circuit issued a decision permitting a local school district to discipline a high school student for comments posted on the student’s blog.1 Doninger v. Niehoff2 is one of the most recent decisions demonstrating the application of Supreme Court language governing on-campus student speech principles to an incident of off-campus, internet-based speech. These cyberspeech cases present a new challenge for courts attempting to apply on-campus First Amendment standards to speech created on the internet and entirely off-campus. While on-campus student speech rights have been carved out in several United States Supreme Court decisions,3 legal scholars have criticized courts for eroding First Amendment rights when applying these standards to cyberspeech occurring entirely off-campus.4 Some First Amendment advocates insist that schools have no authority to prohibit conduct occurring in a student’s home,5 but this position threatens to deprive school administrators of the ability to exercise some measure of authority with which to maintain the discipline and structure of the educational process.

This article argues that Doninger v. Niehoff6 was incorrectly decided because the Second Circuit failed to properly apply student speech standards. First Amendment rights must be balanced with a school administrator’s ability to maintain discipline in the educational process. While it is necessary for the public school system to maintain a degree of authority to ensure the efficiency of the education system and promote the safety of its students and employees, it is also essential that this goal does not infringe on the First Amendment rights that Americans believe to be fundamental to the protection of their democratic

---

1. Doninger v. Niehoff, 527 F.3d 41 (2d Cir. 2008).
2. Id. at 43.
6. 527 F.3d 41 (2d Cir. 2008).
nation. The Second Circuit’s decision threatens to erode these rights because the court inappropriately applied a standard that has worked to balance these competing goals even in the context of off-campus student cyberspeech. The current standard effectively balances these competing goals, and, consequently, judges who consider future controversies should disregard the Second Circuit’s liberal application of this standard.

In order to analyze these issues, Section II explores the relevant background for student speech cases. Specifically, this section examines First Amendment issues and United States Supreme Court cases that have set the standard for on-campus student speech. Courts faced with off-campus cyberspeech cases have used the standards set for on-campus speech in their analyses. Section III examines several off-campus cyberspeech cases that demonstrate the application of these on-campus precedents to off-campus cyberspeech cases. Section IV recites the relevant facts and holdings of Doninger. Section V analyzes the Doninger decision with respect to the facts at issue, the test implemented by the court, and the reasoning employed in similar cases. Section V also argues that the courts are equipped to uphold First Amendment rights and school discipline under the current case law, but the Second Circuit failed to do so in Doninger.

II. BACKGROUND LAW

To properly analyze the issue of student discipline for off-campus cyberspeech, it is first important to understand general First Amendment principles and the relevant case law. To date, the Supreme Court has yet to render a decision implementing a specific standard in off-campus cyberspeech cases. As a result, the federal district and circuit courts have been left to apply First Amendment principles and Supreme Court standards for on-campus speech in such cases. The legal analysis employed, in an effort to resolve these cases, derives from either general First Amendment free speech law or Tinker v. Des Moines and its progeny.

9. See Wisniewski, 494 F.3d at 40 (finding that the Supreme Court has yet to speak on the scope of a school’s authority to regulate expression that does not occur on school grounds or at a school-sponsored event).
A. General First Amendment Principles Governing Student Cyberspeech Cases

First Amendment law provides the basis for relevant analysis of student speech cases.\textsuperscript{11} The First Amendment provides, in part, that "Congress shall make no law . . . abridging the freedom of speech."\textsuperscript{12} While the First Amendment guarantees freedom of speech, it is clear that this right of free speech is not absolute.\textsuperscript{13} Constitutional law permits the imposition of different levels and types of restrictions on speech, depending on the nature of the government property on which the speech is made.\textsuperscript{14} Speech that occurs in public schools is generally subject to reasonable time, place, and manner restrictions, so long as the restrictions are viewpoint neutral.\textsuperscript{15} While this article discusses discipline for student speech, such as internet blogging that occurs entirely off-campus, the regulation of on-campus speech remains relevant to the inquiry because courts addressing this topic have applied tests that were created by the United States Supreme Court in traditional on-campus speech cases.\textsuperscript{16}

Because this article addresses student speech, there are several other aspects of First Amendment law that are implicated. For example, courts charged with considering the reasonableness of a particular regulation limiting student expression have recognized that the age and maturity of the speaker are important factors.\textsuperscript{17} Generally, the rights of school-aged children are not co-extensive with the rights of an adult, and therefore, younger students may be subject to greater restraints on speech than more mature students.\textsuperscript{18} Finally, most courts recognize that the priority of public schools is to instruct its pupils, thus it follows that freedom of expression issues may take a backseat to legitimate educational concerns.\textsuperscript{19}

Before exploring the Supreme Court standards for student speech cases, it is important to recognize a few other First Amendment themes that are relevant in this context. With respect to First Amendment rights on the internet, the Supreme Court has provided that cyberspeech is afforded the highest level of

\textsuperscript{11} Sarah O. Cronan, \textit{Grounding Cyberspeech: Public School's Authority to Discipline Students for Internet Activity}, 97 KY. L.J. 149, 153 (2008).
\textsuperscript{12} U.S. CONST. amend. I.
\textsuperscript{14} Perry Educ. Ass'n v. Perry Local Educators' Ass'n, 460 U.S. 37, 45 (1983).
\textsuperscript{15} See id.
\textsuperscript{16} See Wisniewski v. Bd. of Educ. of Weedsport Cent. Sch. Dist., 494 F.3d 34, 40 (2d Cir. 2007), \textit{cert denied}, 128 S. Ct. 1741 (2008) (finding that a student may be disciplined for expressive conduct, even conduct occurring off school grounds, when this conduct under the \textit{Tinker} standard would foreseeably create a risk of substantial disruption within the school environment).
\textsuperscript{17} Morse v. Frederick, 127 S. Ct. 2618, 2622 (2007).
\textsuperscript{18} Id.
\textsuperscript{19} See Settle v. Dickson County Sch. Bd., 53 F.3d 152, 155 (6th Cir. 1995) ("Where learning is the focus, as in the classroom, student speech may be even more circumscribed than in the school newspaper or other open forum.").
protection available in the constitution. After Reno, it is clear that the internet is a highly protected medium through which to communicate speech.

Another important concept is that “true threats” are not protected by the First Amendment. While this concept has been defined differently in case law, a traditional definition of whether a “true threat” exists is “if a reasonable person would foresee that an objective rational recipient of the statement would interpret its language to constitute a serious expression.” One commentator has noted, that in the school context, courts have employed a multi-factor balancing test such as: “the reaction of the listeners to the threat; whether the threat was conditional; whether the speaker communicated the threat directly to the victim; whether the speaker had a history of making threats against the victim; and whether the recipient had a reason to believe that the speaker had violent tendencies.” Therefore, depending on the type of threat made, courts have prohibited such speech only after weighing First Amendment rights against the likelihood of harm.

B. United States Supreme Court Standards Governing On-Campus Student Speech

The United States Supreme Court has identified four distinct issues to address when deciding cases involving the discipline of student speech. These issues are (1) the extent to which the speech poses a substantial threat of disruption to the educational process; (2) whether the speech is vulgar or lewd; (3) whether the speech appears to be school sponsored and would be considered contrary to the ideals and mission of the school; and (4) speech viewed as promoting illegal drug use.

1. Tinker v. Des Moines Independent Community School District

Tinker is the landmark case in student speech analyses, and it demonstrates the first of the four issues. In Tinker, three students planned on wearing black
armbands to school to demonstrate their opposition to the Vietnam War.\textsuperscript{31} School officials learned of the plan and adopted a policy prohibiting armbands in the school.\textsuperscript{32} The three students wore the armbands despite this policy and were subsequently suspended.\textsuperscript{33}

The lower courts upheld the school administrator’s right to discipline the students, and, after a series of appeals, the case appeared before the United States Supreme Court.\textsuperscript{34} The district court concluded that the discipline was reasonable because it was based upon a fear that the armbands would create a disturbance in the school environment.\textsuperscript{35} The United States Supreme Court acknowledged that there is a careful balancing act in order to both protect First Amendment rights and the need to control disruptive conduct in the schools.\textsuperscript{36} The Court clarified its stance on student rights by insisting that students do not “shed their constitutional rights to freedom of speech or expression at the schoolhouse gate.”\textsuperscript{37} It is with this mindset that the Court addressed the First Amendment claim in \textit{Tinker}.\textsuperscript{38}

After reviewing the facts, the Court determined that the students did not pose a risk of “substantial disruption of or material interference with school activities . . . .”\textsuperscript{39} The Court further opined that the fact that the school administrator’s policy did not prohibit the wearing of all symbols was also relevant to the discussion.\textsuperscript{40} The Court noted that, “the prohibition of expression of one particular opinion, at least without evidence that is necessary to avoid material and substantial interference with schoolwork or discipline, is not constitutionally permissible.”\textsuperscript{41} The Court’s holding essentially proscribed schools from placing general viewpoint restrictions on speech.\textsuperscript{42} More importantly, however, it provided school administrators the means to restrict speech that poses a risk of substantial disruption with the work or discipline of the school.\textsuperscript{43}

The \textit{Tinker} case provides much of the framework applied to student First Amendment cases that have followed.\textsuperscript{44} It is noted for championing the right of student expression and speech in schools, yet it also provides the test for

\textsuperscript{32} \textit{Id.}
\textsuperscript{33} \textit{Id.}
\textsuperscript{34} \textit{Id.} at 505.
\textsuperscript{35} \textit{Id.}
\textsuperscript{36} \textit{Id.} at 506-07.
\textsuperscript{37} \textit{Tinker}, 393 U.S. at 506.
\textsuperscript{38} \textit{Id.}
\textsuperscript{39} \textit{Id.} at 514.
\textsuperscript{40} \textit{Id.} at 510.
\textsuperscript{41} \textit{Id.} at 511.
\textsuperscript{42} \textit{See id.}
\textsuperscript{43} \textit{Tinker}, 393 U.S. at 511.
\textsuperscript{44} \textit{See, e.g.}, Bethel Sch. Dist. No. 403 v. Fraser, 478 U.S. 675, 680 (1986); Hazelwood Sch. Dist. v. Kuhlmeier, 484 U.S. 260, 266 (1988); Morse v. Frederick, 127 S. Ct. 2618, 2622 (2007).
permitting discipline of student speech.\textsuperscript{45} After \textit{Tinker}, it was constitutionally permissible for a school to discipline student speech in an effort to avoid a substantial disruption or material interference with schoolwork or discipline.\textsuperscript{46} When \textit{Tinker} was decided in 1969, the internet was no more real than a science fiction novel, yet the \textit{Tinker} standard for on-campus speech continues to be applied to off-campus cyberspeech today.\textsuperscript{47} Before exploring this in more detail, it is necessary to discuss the other Supreme Court standards for student speech carved out in the years following \textit{Tinker}.

2. \textit{Bethel School District No. 403 v. Fraser}

The second Supreme Court standard governing student speech concerns discipline for lewd and vulgar comments.\textsuperscript{48} In \textit{Fraser}, a high school student delivered a speech at a school assembly where teachers, administrators, and students were present.\textsuperscript{49} The purpose of the speech was to nominate a fellow student for an elective office with the student council.\textsuperscript{50} Throughout the speech, the student referred to the candidate by using several graphic sexual metaphors.\textsuperscript{51} The administration determined that the speaker had violated a school code of conduct prohibiting obscene language.\textsuperscript{52} The school suspended the student for three days and prohibited him from speaking at the school’s graduation ceremonies.\textsuperscript{53}

The suspended student subsequently filed suit against the school alleging a violation of his First Amendment rights to freedom of speech.\textsuperscript{54} Both the United States District Court for the Western District of Washington and the United States Court of Appeals for the Ninth Circuit found that the school board had erred in disciplining the student.\textsuperscript{55} Essentially, both courts found that the speech was protected under \textit{Tinker}, in that the comments did not pose a threat of disruption to the educational process.\textsuperscript{56}

The United States Supreme Court, on the other hand, found that it is a valid function of public schools to prohibit the use of vulgar and offensive language in public discourse on school grounds.\textsuperscript{57} The Court further explained that school

\textsuperscript{45} \textit{Tinker}, 393 U.S. at 512-13.
\textsuperscript{46} \textit{Id.} at 513.
\textsuperscript{47} See Doninger v. Niehoff, 527 F.3d 41, 50-51 (2d Cir. 2008) (upholding the discipline imposed on a student for off-campus cyberspeech under the \textit{Tinker} foreseeable risk of substantial disruption standard).
\textsuperscript{48} \textit{Fraser}, 478 U.S. at 685.
\textsuperscript{49} \textit{Id.} at 677.
\textsuperscript{50} \textit{Id.}
\textsuperscript{51} \textit{Id.} at 678.
\textsuperscript{52} \textit{Id.}
\textsuperscript{53} \textit{Id.}
\textsuperscript{54} \textit{Fraser}, 478 U.S. at 679.
\textsuperscript{55} \textit{Id.}
\textsuperscript{56} \textit{Id.} at 679-80.
\textsuperscript{57} \textit{Id.} at 683.
authorities have the duty to act *in loco parentis*, to protect students, particularly the younger students, from exposure to sexually explicit and vulgar speech.\textsuperscript{58} The Court distinguished *Fraser* from *Tinker* by noting that the penalties imposed in this case were unrelated to any political viewpoint.\textsuperscript{59} The Court held that the school district acted within its authority in imposing sanctions upon Fraser in response to his offensive speech and that the First Amendment did not prevent a school from determining that vulgar and lewd speech is contrary to the basic educational mission.\textsuperscript{60}


The third Supreme Court standard concerns school speech that may be viewed as impeding the school’s ability to carry out its educational mission.\textsuperscript{61} This standard derives from a case in which former members of the Hazelwood school-sponsored student newspaper filed suit against the school district claiming that their First Amendment Rights were violated when school officials forced the students to remove articles discussing teenage pregnancy and divorce from the newspaper.\textsuperscript{62} Among other issues, the school administrators claimed they were concerned that the identity of the pregnant students would be exposed, and that the article would suggest that the school endorsed teenage pregnancy.\textsuperscript{63} The Court recognized that the *Kuhlmeier* case called upon it to decide a much different question than was proposed in *Tinker*.\textsuperscript{64} While *Tinker* involved the school’s ability to silence a student’s personal expression on school premises, *Kuhlmeier* considered the school’s authority over speech that members of the public might reasonably perceive to be accepted and promoted by the school.\textsuperscript{65} The Court found that school administrators are authorized to exhibit greater control over school-sponsored student expression.\textsuperscript{66} Accordingly, the Court adopted a new test for cases arising in this area.\textsuperscript{67} It held that, “educators do not offend the First Amendment by exercising editorial control over the style and content of student speech in school-sponsored expressive activities so long as their actions are reasonably related to legitimate pedagogical concerns.”\textsuperscript{68} Therefore, the Court laid down a third rule for student speech restrictions holding that school administrators could restrict speech that appeared to be

\textsuperscript{58} Id. at 684.
\textsuperscript{59} Id. at 685.
\textsuperscript{60} *Fraser*, 478 U.S. at 685.
\textsuperscript{62} Id. at 264.
\textsuperscript{63} Id. at 263.
\textsuperscript{64} Id. at 270-71.
\textsuperscript{65} Id. at 271.
\textsuperscript{66} Id.
\textsuperscript{67} Kuhlmeier, 484 U.S. at 273.
\textsuperscript{68} Id.
school sponsored when the restrictions are reasonably related to legitimate pedagogical concerns.69

4. Morse v. Frederick

The final case in this line of Supreme Court rulings concerns speech that promotes the use of illegal drugs.70 In Morse v. Frederick, a high school student alleged that his First Amendment rights had been violated by the imposition of a ten-day suspension for holding a banner on which was written “BONGHiTS 4 JESUS” at an off-campus school sponsored event.71 The Court held that the school’s principal did not violate the student’s First Amendment rights by confiscating the banner that promoted illegal drug use.72 This standard is greatly limited in comparison to the previous three because its holding is narrowly confined to illegalities; and as Justice Alito noted in his concurring opinion, even more narrowly confined to speech promoting illegal drug use.73 Still, this case represents the fourth United States Supreme Court standard encompassing the guiding case law on student First Amendment issues.

III. Student Cyberspeech Cases Prior to Doninger

The preceding section detailed the United States Supreme Court standards for on-campus student speech. Since the advent of the internet, First Amendment concerns for off-campus speech have become increasingly relevant. The following cases illustrate approaches taken by various lower courts faced with First Amendment claims deriving from off-campus cyberspeech.

With respect to off-campus cyberspeech, the Tinker analysis is employed in the majority of cases.74 Morse will generally not apply to these cases as they do not take place at school-sponsored events.75 Similarly, it would be rare for a student’s blog or MySpace page to reasonably appear to be expressing a school sponsored ideology as to place the facts within the context of Kuhlmeier.76 Finally, courts have typically declined to extend the Fraser standard for vulgar and lewd speech occurring entirely off-campus.77

It should be noted, however, that at least one court has preemptively circumvented the threshold question of whether the speech occurred on or off of

---

69. Id.
70. Morse v. Frederick, 127 S. Ct. 2618, 2622 (2007).
71. Id.
72. Id. at 2629.
73. Id. at 2636 (Alito, J., concurring).
74. See discussion infra Part III.A-B.
75. See Morse v. Frederick, 127 S.Ct. 2618 (2007).
77. See, e.g., Doninger v. Nichoff, 527 F.3d 41, 49 (2d Cir. 2008) (finding that though the student’s conduct clearly would have satisfied the Fraser standard if it had occurred on-campus, it is unclear that Fraser applies to off-campus student speech, and therefore, declining to analyze the issue under Fraser).
school grounds. By doing so, the judiciary is more comfortable with classifying the speech under any of the Supreme Court standards. Courts have established that student cyberspeech occurred on-campus when the speech was either written on a school computer, accessed on a school computer, or brought to campus by another student. Because this article attempts to provide a commentary on truly off-campus student speech, this article will not focus upon this issue. However, the court’s analysis of whether a substantial disruption existed in these cases remains relevant in determining what is, and what is not, a substantial disruption.

The subsequent body of case law demonstrates the typical analysis employed in student Internet cases prior to Doninger. These cases illustrate both facts that have led courts to protect student cyberspeech under Tinker and facts that have led courts to uphold a school’s decision to discipline student cyberspeech under the same standard.

A. Cyberspeech Satisfying the Tinker Standard

1. Wisniewski v. Board of Education of Weedsport Central School District

In Wisniewski, an eight grade student brought a First Amendment claim against his school district after school administrators suspended the student for an internet icon he created. The student used AOL Instant Messaging software to create a “buddy icon” on his parents’ home computer. The icon in question could be viewed by the students online “buddies.” The icon depicted a pistol firing a bullet at a person’s head and blood splattering from the head. Beneath the drawing appeared a message reading “Kill Mr. VanderMolen,” who was the student’s English teacher. A fellow student viewed the icon and brought it to the attention of the targeted teacher. The teacher notified school officials who initially suspended the student for five days, and following a hearing with the superintendent, suspended the student for the entire semester. Meanwhile, the

78. See J.S. ex rel H.S. v. Bethlehem Area Sch. Dist., 807 A.2d 847, 865 (Pa. 2002) (finding that because the student accessed the webpage at issue from a school computer, and displayed the webpage to students at school, the court could properly address the student’s conduct under any of the First Amendment standards established by the United States Supreme Court for on-campus speech).
79. Id.
80. See id.
81. See Doninger, 527 F.3d at 49-50.
83. Id. at 35.
84. Id.
85. Id. at 36.
86. Id.
87. Id.
88. Wisniewski, 494 F.3d at 36.
teacher was permitted to stop instructing the class in which the student was enrolled.\textsuperscript{89}

The court found that the even though the icon was created away from school property, the student was not necessarily insulated from school discipline.\textsuperscript{90} The court explained that under the \textit{Tinker} standard, it was reasonably foreseeable that the icon would come to the attention of school authorities and to the teacher depicted in the icon.\textsuperscript{91} The court further reasoned that it was reasonably foreseeable due to the fact that the student distributed the icon to at least fifteen “buddies,” some of which were classmates, over a three-week period.\textsuperscript{92} The court determined that the icon, once made known to the teacher and school officials, would foreseeably create a risk of substantial disruption within the school environment.\textsuperscript{93} Therefore, the court held that school could properly discipline the student for the creation and use of the buddy icon.\textsuperscript{94}

2. \textit{J.S. ex rel H.S. v. Bethlehem Area School District}

In \textit{J.S. ex rel H.S. v. Bethlehem Area School District}, the court upheld the discipline of a student for content posted on the student’s personal website.\textsuperscript{95} The website, named “Teachers Sux,” was created on the student’s home computer during his personal time.\textsuperscript{96} The website contained a number of web pages that made derogatory and threatening remarks about the student’s algebra teacher and middle school principal.\textsuperscript{97} Among other comments, the website included allegations that the teacher had sexual relations with a principal, contained a section called “Why should she die?,” featured a depiction of the teacher’s severed head dripping blood, and requested that visitors to the site contribute twenty dollars to “help pay for the hit man.”\textsuperscript{98}

One teacher learned of the site from school administrators and took the threats seriously.\textsuperscript{99} He informed the algebra teacher, and she applied for and was granted medical leave after suffering stress and anxiety as a result of the student’s website.\textsuperscript{100} As a result of the teacher’s inability to return to work, three substitute teachers were required to fulfill her obligations.\textsuperscript{101}

\begin{itemize}
\item \textsuperscript{89} \textit{Id.}
\item \textsuperscript{90} \textit{Id.} at 39-40.
\item \textsuperscript{91} \textit{Id.}
\item \textsuperscript{92} \textit{Id.}
\item \textsuperscript{93} \textit{Id.} at 40.
\item \textsuperscript{94} \textit{Wisniewski}, 494 F.3d at 40.
\item \textsuperscript{95} \textit{J.S. ex rel H.S. v. Bethlehem Area School District}, 807 A.2d 847, 869 (Pa. 2002).
\item \textsuperscript{96} \textit{Id.} at 850-51.
\item \textsuperscript{97} \textit{Id.} at 851.
\item \textsuperscript{98} \textit{Id.}
\item \textsuperscript{99} \textit{Id.} at 852.
\item \textsuperscript{100} \textit{Id.}
\item \textsuperscript{101} \textit{Bethlehem}, 807 A.2d at 852.
\end{itemize}
The Bethlehem court, relying on Tinker, determined that an actual, material disruption of the school environment would permit the school district to take action, and school officials did not have to wait for possible harm or material disruption to come to pass before taking appropriate steps. The court found that the school had sustained its burden of establishing that the student’s speech materially disrupted class work under the Tinker standard, specifically, by showing the emotional and physical impacts on the teacher who was forced to miss time at school.

It is important to note in Bethlehem, the court also analyzed whether there was a sufficient nexus between the website and the school campus to consider the speech as occurring on campus. The record indicated that the student accessed the website at school and displayed the site to other students while at school. Furthermore, the website was aimed at the specific audience of students and school employees who were the subjects of the site. As such, the court determined the website to be on-campus speech, and thus it was free to apply any of the Supreme Court standards for on-campus speech. In doing so, the court found that the school’s discipline could be upheld under the Fraser standard governing lewd and vulgar speech. The court, however, refrained from upholding the decision under Fraser as it noted that questions existed as to the applicability of Fraser to the facts in this case. Alternatively, the court examined cyberspeech cases that effectively applied Tinker, and decided to uphold the school’s decision under the substantial disruption standard.

B. Cyberspeech Failing to Satisfy the Tinker Standard


In Flaherty, school administrators disciplined a student for posting disparaging remarks on the internet about a teacher and the high school men’s volleyball team. These messages were posted on a website message board that

102. Id. at 861.
103. Id. at 869.
104. This article does not attempt to analyze decisions that have considered cyberspeech to have a sufficient nexus with the school campus as to be considered on-campus speech. However, the decision in Bethlehem remains relevant to this article because the discipline was upheld under Tinker, which is the standard employed in off-campus cyberspeech cases. Therefore, the facts in Bethlehem remain relevant in determining the type of off-campus cyberspeech that will be subject to discipline under the Tinker standard.
105. Bethlehem, 807 A.2d at 865.
106. Id.
107. Id.
108. Id.
109. Id. at 868.
110. Id.
111. Bethlehem, 807 A.2d at 868-69.
was not affiliated with the school district. The student posted several messages from a personal computer at his parents’ home and one message from a computer at the school. The school disciplined the student based on provisions in the district’s student handbook that allowed discipline for verbal or written abuse directed towards a school employee and other similar provisions.

The Flaherty court overturned the school district’s discipline of the student. The court followed the United States Supreme Court guidelines set forth in Tinker and found no evidence that the student’s internet messages substantially disrupted the normal operations of the school district or the rights of other students. The court further reasoned that, both inside and outside of the school context, the mere fact that someone might take offense at the content of speech is not sufficient justification for prohibiting it.

2. Mahaffey ex rel Mahaffey v. Aldrich

In Mahaffey, parents of a student filed suit against the school district as a result of the expulsion of their son for content posted on a website. The website in question was created by another student and was titled “Satan’s Web Page.” The website included a listing of “people I wish would die,” “people that are cool,” “movies that rock,” and other opinions expressed by the student. It also contained a section called, “SATAN’S MISSION FOR YOU THIS WEEK,” which encouraged the reader to stab someone for no reason, and then included a disclaimer reading, “NOW THAT YOU’VE READ MY WEB PAGE PLEASE DON’T GO KILLING PEOPLE AND STUFF THEN BLAMING IT ON ME. OK?” The school district suspended and recommended expulsion of the student under provisions of the student code of conduct pertaining to “behavior dangerous to self and others,” “internet violations,” and “intimidations and threats.”

The Mahaffey court held that “[the district] may only punish the [student] for his speech on the website if that speech ‘substantially interfere[d] with the work of the school or impinge[d] upon the rights of other students.’” Ruling in

113. Id.
114. Id.
115. Id. at 701.
116. Id. at 705.
117. Id.
118. Flaherty, 247 F. Supp. 2d at 705.
120. Id.
121. Id. at 782.
122. Id.
123. Id. at 782-83.
124. Id. at 784 (citing Tinker v. Des Moines Indep. Cmty. Sch. Dist., 393 U.S. 503, 509 (1969)).
favor of the student, the Mahaffey court found no evidence that the website interfered with or encouraged the interference with the normal operations of the school district or that any other student’s rights were impinged. Additionally, the court determined that the website did not constitute a “true threat” as to limit the content of the website under the First Amendment. Instead, the court considered the website to be a crude attempt at humor that was not intended to be communicated to, or threaten, anyone else.


In Killion, a student created a disparaging “top ten list” about the high school athletic director. The student created the list on his home computer and e-mailed it to several of his friends. The list appeared on school grounds, though the student had nothing to do with its presence on school property. The administration suspended the student for ten days; subsequently, the student filed suit seeking a preliminary injunction against the suspension for violation of his First and Fourteenth Amendment rights.

The court granted the student’s motion for summary judgment on the grounds that his First Amendment rights were violated by the school district when it suspended him. The court found that although the document contained vulgar and lewd language, the school could not suspend the student for this content when it was created off school grounds. The court further reasoned that, because the document was created off-campus, and the student did not bring the list to school, the school could only discipline the student if it could show a substantial likelihood of disruption. The court found that there was “no evidence that teachers were incapable of teaching or controlling their classes because of the . . . list. Indeed, the list was on school grounds for several days before the administration became aware of its existence, and at least one week passed before the defendants took any action.”

The court also determined that the speech was not threatening. The speech, while insulting, did not cause any faculty member to require a leave of absence, as was the case in J.S. v. Bethlehem. The court insisted that disliking...

125. Mahaffey, 236 F. Supp. 2d at 786.
126. Id.
127. Id.
129. Id.
130. Id. at 448-49.
131. Id. at 449.
132. Id. at 455, 458.
133. Id. at 457.
135. Id. at 455.
136. Id.
137. Id. See case cited supra notes 99-101.
the speech or the mere desire to avoid discomfort was not an acceptable ground for punishment under *Tinker*.

The court found that, “[a]t best, defendants have offered evidence that . . . [the athletic director] was upset and had a hard time doing his job . . . However, these events do not rise to the level of substantial disruption, and do not support an expectation of disruption defense.”

Although, the speech undoubtedly targeted people at the school, the court chose not to uphold the punishment without evidence of threats or actual disruption.

4. *Latour v. Riverside Beaver School District*

In *Latour*, a student was suspended and subsequently expelled for two years as a result of four rap songs, recorded over a two-year period, that discussed other students and contained violent language. The student wrote and recorded all of the songs in his own home, and never brought the songs in written or recorded form to school. The student filed a motion for a preliminary injunction, seeking to enjoin the school from enforcing the expulsion. The school district argued that its disciplinary measures were constitutional because the rap songs were either “true threats” or created a likelihood of substantial disruption.

The court dismissed the district’s argument that the songs were true threats. The court reasoned that the students mentioned in the songs did not personally feel threatened and the student responsible for the songs had no history of violence. Furthermore, the court determined that although the rap songs contained violent imagery, they were written in the mold of true rap songs that had no intention of actual violence. Finally, the court also noted that the songs were not directly communicated to the individuals targeted in the lyrics.

The court rejected the district’s second argument that the songs created a substantial disruption in the school. The court found that any disruption that had occurred was a result of the discipline conferred and not of the rap songs themselves. Therefore, the court found that the student was likely to prevail on the merits of the case.

---

138. *Id.*
139. *Id.* at 455-56.
142. *Id.*
143. *Id.*
144. *Id.*
145. *Id.* at *2.
146. *Id.*
148. *Id.*
149. *Id.*
150. *Id.*
151. *Id.*
5. *Beussink ex rel. Beussink v. Woodland R-IV School District*

In *Beussink*, a student requested a preliminary injunction claiming that his First Amendment rights were violated when he was suspended for posting a homepage on the internet that was critical of his high school. The high school senior created a homepage that could be accessed by other internet users. The homepage contained vulgar language expressing the student’s negative opinions pertaining to the teachers and principal of the school. It also contained a link to the school’s homepage and encouraged readers to contact the school to communicate their opinions with the principal. The administration learned of the homepage after another student showed the website to the computer teacher whereupon the principal issued an immediate five-day suspension, which was subsequently increased to ten days.

In considering the students motion for a preliminary injunction, the court examined the *Tinker* standard to determine that speech may be limited when there is a fear of disruption, but that fear must be reasonable. The court, however, noted that the principal was merely upset by the content of the homepage reasoning that, “[d]isliking or being upset by the content of a student’s speech is not an acceptable justification for limiting student speech under *Tinker*.” Therefore, the court found that the student was likely to succeed on the merits of his claim.

The foregoing cases illustrate a variety of off-campus student activity for which students have been disciplined. Each of the preceding cases was examined under the *Tinker* standard to determine whether the discipline violated the First Amendment rights of the students in question. These cases provide a general background of the law governing student cyberspeech issues while the following sections explore the outcome of *Doninger v. Niehoff* within this context.

---

153. *Id.*
154. *Id.*
155. *Id.*
156. *Id.*
157. *Id.* at 1178-79.
159. *Id.*
160. *Id.*
161. *Id.*
162. 527 F.3d 41 (2d Cir. 2008).
IV. RELEVANT FACTS AND HOLDING IN DONINGER

Avery Doninger attended Lewis Mills High School (hereinafter “LMHS”) in Burlington, Connecticut as a junior during the 2006-2007 school year. Doninger served as a member of the Student Council and was also the Junior Class Secretary. In the spring of 2007, LMHS planned a battle-of-the-bands concert known as “Jamfest” and the Student Council helped the administration plan the concert. During the 2007 school year, LMHS was finishing construction of a new auditorium. Delays in the construction caused the school administration to postpone “Jamfest” on two occasions. Following the second postponement, the concert was rescheduled to take place on Saturday, April 28. Prior to this date, however, the teacher in charge of the sound and lighting equipment indicated that he would be unable to attend the concert on the 28th. As a result, Doninger and her fellow Student Council members met with the administration and proposed that the school hire a professional to operate the equipment for the concert. The school refused to comply with the Council’s request and insisted that unless the concert could be held in the cafeteria, the concert would, again, be postponed. The administration’s decision frustrated the students who believed that the cafeteria was an unsuitable alternative and feared that there were few dates remaining in the school year on which to host the event.

As a result of the administration’s decision, Doninger and three other Student Council members sent a mass email at the school’s computer lab to the community explaining the situation and urging people to contact the district superintendent and request that the school host “Jamfest” on the 28th of April. The student’s email elicited a large response from the community, and both the principal and superintendent were inundated with emails and phone calls. Principal Karissa Niehoff subsequently met with Doninger to discuss the email. She expressed her disappointment with the Student Council members explaining that they should have approached her rather than resorting to a mass
Niehoff further stated that she was open to rescheduling the event and requested that Doninger send out a corrective email to explain the situation.

That evening, Doninger posted a message on livejournal.com, a publicly accessible blog that was unaffiliated with the school. In the message, Doninger stated that “Jamfest” was canceled “due to douchebags in central office.” Doninger also reproduced the email that was sent out requesting community support and an email that her mother had sent to the superintendent. Finally, she encouraged people to contact the superintendent to “piss her off more.”

The administration, unaware of Doninger’s blog, met with the students to discuss the “Jamfest” issue on April 28. In the meeting the administration asserted that resorting to a mass email was an inappropriate action for class officers attempting to resolve a conflict with the administration. The administration and Council were able to successfully reschedule the concert for June 8.

In May, however, the superintendent’s adult son found Doninger’s blog on the internet. The superintendent notified Principal Niehoff of the blog, and she concluded that Doninger’s conduct was unbecoming of a class officer. As a result, Niehoff determined that Doninger should be prohibited from running for Senior Class Secretary. Niehoff explained that the decision was based on three principles: 1) Doninger’s failure to exercise the proper manner of expressing her disagreement with the administration’s policy; 2) the vulgar language and inaccurate information contained in the internet post; and 3) Doninger’s encouragement of others to contact the superintendent to “piss her off more.”

Despite the fact that Doninger’s name was stricken from the class secretary ballot, she received a majority of the votes as a write-in. Niehoff, however, awarded the position to the second place finisher. Doninger’s mother subsequently filed suit, on her behalf, in Connecticut Superior Court. The claim alleged that the Niehoff and the superintendent had violated Doninger’s

176. Id. at 45.
177. Id.
178. Id.
179. Id.
180. Doninger, 527 F.3d at 45.
181. Id.
182. Id. at 46.
183. Id.
184. Id.
185. Id.
186. Doninger, 527 F.3d at 46.
187. Id.
188. Id.
189. Id.
190. Id.
First Amendment rights, under the United States and Connecticut Constitutions, and her due process and equal protection rights under the Fourteenth Amendment.\footnote{191. \textit{Id.} at 47.} The claim additionally asserted a cause of action for intentional infliction of emotional distress and prayed for an injunction requiring the school to hold a new election with Doninger on the ballot.\footnote{192. \textit{Doninger}, 527 F.3d at 47.}

The defendants removed the lawsuit to the District Court of Connecticut.\footnote{193. \textit{Id.}} Doninger subsequently filed a motion for a preliminary injunction, which the district court denied by concluding that based on the facts and testimony of witnesses, Doninger was unable to demonstrate a sufficient likelihood of success on the merits.\footnote{194. \textit{Id.}} Doninger appealed the district court’s decision to the Federal Court of Appeals for the Second Circuit.\footnote{195. \textit{Id.}}

The Second Circuit held that the district court did not abuse its discretion in finding that Doninger did not show a substantial likelihood of success on the merits of the First Amendment claim.\footnote{196. \textit{Id.}} In doing so, the court used the Supreme Court’s standards for regulating on-campus speech and relied on Second Circuit precedent.\footnote{197. \textit{Id. at} 48-51.} The court began by laying out the Supreme Court guidelines finding that student speech can be regulated: when it poses a threat of materially and substantially disrupting the work and discipline of the school, when the student uses vulgar or offensive speech, or when regulation over a school sponsored activity is related to legitimate pedagogical concerns.\footnote{198. \textit{Doninger}, 527 F.3d at 48.}

The Second Circuit then noted that, “[t]he Supreme Court has yet to speak on the scope of a school’s authority to regulate expression that...does not occur on school grounds or at a school-sponsored event.”\footnote{199. \textit{Id.}} The Second Circuit instead turned to a case it had decided in 2007.\footnote{200. 494 F.3d 34 (2d Cir. 2007), \textit{cert. denied}, 128 S. Ct. 1741 (2008).} In \textit{Wisniewski v. Board of Education},\footnote{201. \textit{Doninger}, 527 F.3d at 48. \textit{See supra} notes 82-94.} the Second Circuit determined that a student may be disciplined for expressive conduct that occurs off school grounds, when this conduct would create a foreseeable risk of a substantial disruption within the school environment.\footnote{202. \textit{Id. at} 49.}

After establishing that Doninger’s speech was subject to potential discipline, the court then examined the speech within the context of related cases.\footnote{203. \textit{Id. at} 49.} The court found that the internet posting, referring to the school administrators as “douchebags” and encouraging people to contact the superintendent to “piss her
off more,” satisfied the Fraser standards for vulgar and offensive language.\textsuperscript{204} The court then noted, however, that it was unclear whether this Fraser principle applied to off-campus speech.\textsuperscript{208} While the court denied that it extended the Fraser standard in the Wisniewski decision, it affirmed that its holding in Wisniewski applied the Tinker substantial disruption standard to off-campus speech.\textsuperscript{206} Using the Tinker analysis, the court found that it was reasonably foreseeable that Doninger’s posting would reach school grounds.\textsuperscript{207} The court accepted the notion that the posting was designed to illicit student response, and that the posting did in fact eventually reach school administrators.\textsuperscript{208}

After determining that it was foreseeable that the posting would reach school grounds, the court examined whether the posting created a foreseeable risk of substantial disruption within the school environment.\textsuperscript{209} The court focused on three factors in determining that the internet posting created a risk of substantial disruption.\textsuperscript{210} First, it explained that the vulgar language used to elicit a response could create disruption.\textsuperscript{211} Next, it found that the post contained false or misleading information concerning the cancellation of “Jamfest” and that this false information “riled up” the students.\textsuperscript{212} Meetings with the students and school administrators were required to correct the false information.\textsuperscript{213} The court additionally refuted the idea that Tinker requires a showing of actual disruption in the school environment.\textsuperscript{214} Instead, the court insisted that school officials have a duty to prevent foreseeable disruption.\textsuperscript{215}

Finally, the court explored the nature of Doninger’s student council position.\textsuperscript{216} It found that her actions were not in line with the standards of behavior the school required for student council members.\textsuperscript{217} In the context of Tinker, the court determined that Doninger’s internet posting materially and substantially disrupted the work and discipline of the school by undermining the values that student government is designed to promote.\textsuperscript{218} Using this analysis, the Second Circuit concluded that the district court did not abuse its discretion in

\textsuperscript{204.} Id.
\textsuperscript{205.} Id.
\textsuperscript{206.} Id. at 50.
\textsuperscript{207.} Id.
\textsuperscript{208.} Doninger, 527 F.3d at 50.
\textsuperscript{209.} Id.
\textsuperscript{210.} Id.
\textsuperscript{211.} Id. at 51.
\textsuperscript{212.} Id.
\textsuperscript{213.} Id.
\textsuperscript{214.} Doninger, 527 F.3d at 51.
\textsuperscript{215.} Id.
\textsuperscript{216.} Id. at 52.
\textsuperscript{217.} Id.
\textsuperscript{218.} Id.
determining that Doninger was unlikely to succeed on the merits to satisfy her motion for a preliminary injunction.219

V. ANALYSIS

The Second Circuit correctly addressed the First Amendment issue created by Doninger’s blog under the foreseeable risk of substantial disruption standard implemented by the United States Supreme Court in Tinker.220 The court, however, incorrectly applied this standard to the facts of the case. The court was unable to rely on evidence that reasonably supported a finding that Doninger’s blog created a foreseeable risk of substantial disruption. Instead, the court cited factors pertaining to the vulgar nature of Doninger’s blog, and the general right of participation in extracurricular activities, to support its weak inference that Doninger’s speech could be disciplined under Tinker.221 In essence, the court justified the discipline because, in the court’s determination, the discipline itself was not offensive in relation to Doninger’s behavior.222 However, by using Tinker to justify its decision, the court incorrectly imposed discipline under the guise of a First Amendment dispute. The court’s willingness to sanction the speech in the blog posting resulted in a misguided extension of the substantial disruption standard.223 The Second Circuit’s decision, therefore, may expose future First Amendment cyberspeech claims to this imprudent expansion of Tinker.

The Second Circuit focused on three factors to incorrectly determine that the school could discipline Doninger without infringing on her First Amendment rights.224 The court first focused on the language Doninger used in her blog, specifically in referring to the central office as “douchebags.”225 The court erroneously concluded that this language evidenced the possibility of creating a substantial disruption.226 It is illogical to conceive that the “vulgar” terms “douchebags,” or “piss her off,” would by themselves be likely to incite high-school-aged students to disrupt the school.227 Furthermore, while Fraser permits school administrators to sanction students for vulgar and lewd speech on campus,228 the Second Circuit dismissed Fraser in its analysis because the standard prohibiting lewd and vulgar speech has traditionally been limited to on-
c Campus speech. Yet, the Second Circuit’s opinion thoroughly discussed the terms used in Doninger’s blog in conjunction with analyzing Doninger’s values and civility. It is indeed confounding that the court would dismiss the legal analysis in Fraser then spend an inordinate amount of time discussing the “crude” language, in an effort to invent a causal link between the language and a chance of potential disruption.

In spite of this legal “two-step,” the question arises as to what disruption can truly be evidenced by using the term “douchebags” on a blog. School administrators are completely out of touch with reality if they believe that this admittedly crude language is anything but commonplace in student discussions of school employees. If anything, the “douchebag” criticism offers evidence that the school’s decision to discipline Doninger was a retaliatory effort aimed at Doninger for the name-calling. In fact, it seems apparent that while referring to a teacher or administrator as a “douchebag” on-campus may fall under the Fraser analysis, it certainly would not substantially disrupt the school even if shouted in the schoolyard. High-school-aged students often communicate with each other using language much more vulgar than “douchebag,” and it cannot be properly deduced that this “P.G.” term would “rile up” a student body of this age. Therefore, it is clear that the “crude” language, in and of itself, did not create a foreseeable risk of substantial disruption.

After incorrectly finding that the language used by Doninger indicated a potential risk of disruption, the court then turned to evidence that Doninger’s posting was “misleading” or “inaccurate” with respect to “Jamfest” being cancelled. While the court focused on the idea that students were “riled up” and the administrators were inundated with emails and phone calls about “Jamfest’s” cancellation, the fact remains that the students believed that the third postponement would leave no suitable dates on which to hold the event. The school’s decision to discipline Doninger, however, must be based on the potential disruption created by students or community members contacting school officials and “divert[ing] [them] from their core educational responsibilities.”

This argument also fails to bring the discipline within the confines of Tinker for two reasons. First, Doninger’s blog was not the primary source of the students’ and community’s discontent. The word was spread around campus
and the community by Doninger and other student council members prior to the blog, when Doninger and others decided to send an email to the community.\textsuperscript{239} The court noted that Doninger and other students met with school administrators on April 25, to discuss the so-called potential disruption.\textsuperscript{240} This meeting occurred weeks before school administrators discovered Doninger’s blog.\textsuperscript{241} The alleged disruption that occurred when community members contacted the superintendent was in response to this email, not to Doninger’s blog.\textsuperscript{242} The meeting with the students properly addressed this problem, and only when Doninger’s blog was discovered weeks later, did the administration choose to discipline Doninger.\textsuperscript{243} Notably, the school administrators did not discipline the other students who were responsible for the initial email. This retroactively imposed punishment provides further evidence that school administrators disciplined Doninger due to their personal discontent with Doninger for referring to the central office as “douchebags.”\textsuperscript{244}

Several cyberspeech cases, however, have expressly held that school officials may not discipline student speech simply because they are insulted by, or dislike the content of, the speech.\textsuperscript{245} The court’s analysis insists that the confusion over “Jamfest” implies that it was foreseeable that school administrators would be “diverted from their core educational responsibilities” in the future to dissipate this confusion.\textsuperscript{246} Again, this “confusion” is attributable to the email, which was sent by the students prior to posting of the targeted message in Doninger’s blog.

The court also failed to bring the discipline within the confines of \textit{Tinker}\textsuperscript{247} for a second reason. The “disruption” created when school administrators were inundated with phone calls about “Jamfest” does not rise to the level of disruption necessary to restrict speech under \textit{Tinker}.\textsuperscript{248} “Jamfest” had already been rescheduled three times and the students believed no suitable dates remained in the academic year to reschedule “Jamfest.”\textsuperscript{249} These facts offer a legitimate reason for students and members of the community to seek answers from school administrators. It can hardly be considered a substantial disruption to meet with a small number of students to discuss discontent that should have been foreseeable to the administration when it delayed “Jamfest” for the third

\begin{itemize}
  \item \textsuperscript{239} Id.
  \item \textsuperscript{240} Id. at 46.
  \item \textsuperscript{241} \textit{Doninger}, 527 F.3d at 46.
  \item \textsuperscript{242} Id.
  \item \textsuperscript{243} Id.
  \item \textsuperscript{244} Id. at 45.
  \item \textsuperscript{246} \textit{Doninger}, 527 F.3d at 51.
  \item \textsuperscript{248} \textit{Doninger}, 527 F.3d at 50.
  \item \textsuperscript{249} Id. at 44.
\end{itemize}
time. Similar to *Latour*, the substantial disruption alleged in *Doninger* was created by the school district’s decisions. In this case, the repeated decisions to delay “Jamfest” created the response that school officials claimed to be a substantial disruption. The students and community responded to the administration’s decision to delay “Jamfest” and not to the message posted in Doninger’s blog. Furthermore, it would appear to be a duty of the superintendent to answer student and community concerns pertaining to a school scheduled event rather than a substantial disruption to the daily tasks of the superintendent.

Additionally, the fact that school administrators may have to respond to student concerns regarding school-sanctioned events can hardly be viewed as the sort of substantial disruption created by the students in similar cyberspeech cases. In *Wisniewski* for example, a teacher had to alter his teaching schedule in response to the student’s icon, which depicted the teacher being shot in the head. In *Bethlehem*, teachers experienced mental and physical ailments resulting from the student’s cyberspeech. One teacher was forced to take medical leave from the school and three substitute teachers had to be employed to handle her classes. These cases demonstrate an actual disruption in the discipline and work of the school. The *Doninger* facts share more similarities with the “top ten list” created in *Killion*, where although school administrators were insulted by the disparaging remarks made in the list, no disruption could be shown to satisfy the *Tinker* standard.

While it is true that school administrators do not have to wait for an actual disruption to take place, the argument in *Doninger* was that the blog created a foreseeable risk of diverting administrators from their core duties. Because the fear of disruption did not stem from Doninger’s blog and the disruption that was feared was in no way substantial, the court’s analysis as to this second factor also fails to evince support for disciplining Doninger for her blog posting.

250. *See id.*
252. *Doninger*, 527 F.3d at 44.
253. *Id.*
256. *Id.*
258. *See, e.g., Lowery v. Euverard, 497 F.3d 584, 596 (6th Cir. 2007), cert. denied, 129 S. Ct. 159 (2009) (“School officials have an affirmative duty to not only ameliorate the harmful effects of disruptions, but to prevent them from happening in the first place.”); see also LaVine v. Blaine Sch. Dist., 527 F.3d 981, 989 (9th Cir. 2001) (“*Tinker* does not require school officials to wait until disruption actually occurs before they may act.”).*
259. *Doninger v. Niehoff, 527 F.3d 41, 51 (2d Cir. 2008).*
260. *See id.*
The third and final section of the court’s reasoning is perhaps the most unpersuasive. The Second Circuit writes, “the district court correctly determined that it is of no small significance that the discipline here related to Avery’s extracurricular role as a student government leader.” The court additionally determined that it had “no occasion to consider whether a different, more serious consequence than disqualification from student office would raise constitutional concerns.” The degree of discipline imposed on Doninger, however, is irrelevant to whether discipline may be imposed as a result of her blog. While students do not have a right to participate in extracurricular activities, they still may not be disciplined for speech that is protected by the First Amendment.

The court relied on multiple facts in the record to rationalize that the discipline was not offensive. But it is imprudent to canvas the analysis of this discipline in the First Amendment context. As the Court noted in Tinker, “In the absence of a specific showing of constitutionally valid reasons to regulate their speech, students are entitled to freedom of expression of their views.” First Amendment jurisprudence is not concerned with the degree of discipline; rather it is concerned with whether the speech is protected. In other words, the degree of discipline should not have impacted the court’s decision as to whether Doninger’s blog created a foreseeable risk of substantial disruption in the school. By qualifying its judgment in terms of the discipline, the court is altering the true First Amendment standard and incorporating a separate constitutional concern of due process. Tinker was concerned about whether student speech could be disciplined in compliance with First Amendment jurisprudence, not whether speech could be disciplined to a certain degree. Therefore, this third analysis works to inappropriately justify the school’s decision to discipline Doninger for her internet posting because, in the court’s view, the school imposed only a minor degree of discipline.

The Second Circuit’s decision in Doninger resembles other student cyberspeech cases in that it instituted the Tinker “foreseeable risk of substantial disruption” standard to determine whether the discipline imposed on Doninger violated her First Amendment rights to free speech. The court, however,
incorrectly applied this standard to the facts of the case. The Second Circuit’s decision set an unfortunate precedent by focusing on the wrong facts. It will likely encourage school administrators to discipline students for off-campus speech which administrators find insulting. Furthermore, the decision may thwart student activism and chill student’s First Amendment rights. School administrators must be provided the opportunity to maintain the discipline of their schools, but sacrificing First Amendment rights in cases such as Doninger creates a worrisome standard. True substantial disruptions created by cyberspeech, like those found in Wisniewski or Bethlehem, may be properly disciplined by administrators under the Tinker standard. Doninger on the other hand, was the victim of a court that was willing to sacrifice First Amendment protections to justify a “reasonable” level of discipline. The Second Circuit’s abuse of the Tinker standard in the Doninger case threatens to alter the Tinker standard that has effectively been applied to off-campus cyberspeech cases in the past.

VI. CONCLUSION

School-aged children will not always have pleasant things to say about their teachers and school administrators. The law should recognize that such behavior was commonplace prior to the internet and will continue to be so in the future. School administrators and judges need to recognize that mere “trash talking” on the internet should not invoke the school administrators’ ability to extend their disciplinary measures into the student’s own home. Impulsive reactions by school administrators who are motivated by personal humiliation due to critical commentary must not be tolerated. The internet has, however, provided a new medium in which to share speech with others at an unprecedented rate. This forum of communication cannot shelter students from disciplinary measures merely because it involves off-campus speech. In reality, students should recognize that, depending on the nature of their internet activity, they may be expressing their ideas to a multitude of people. Thus, while simple insults need not invoke discipline, truly disruptive messages should be deterred.

The Tinker substantial disruption test equips courts with an appropriate framework to decide student cyberspeech cases. While Tinker was decided years before students began blogging or instant messaging each other online, the Court’s standard has proved workable in the cyberspeech context. Clearly, student speech deserves more protection when it occurs off of school grounds in order to avoid chilling First Amendment speech. It is, however, inherently less
foreseeable that internet speech will cause any sort of disruption in the classroom; therefore, the Tinker standard naturally provides students with greater protection in these off-campus cyberspeech cases. Critics of Doninger should recognize that there is a need to prevent truly disruptive speech in public schools. Unfortunately, the Second Circuit failed to properly apply the standards set forth by the United States Supreme Court in Tinker, and therefore, this decision should not represent the state of First Amendment rights for students.