FREE SPEECH CONSEQUENTIALISM

*Erica Goldberg**

Balancing the harms and benefits of speech—what this Article calls "free speech consequentialism"—is pervasive and seemingly unavoidable. Under current doctrine, courts determine if speech can be regulated using various forms of free speech consequentialism, such as weighing whether a particular kind of speech causes harms that outweigh its benefits, or asking whether the government has especially strong reasons for regulating particular kinds of speech. Recent scholarship has increasingly argued for more free speech consequentialism. Scholars maintain that free speech jurisprudence does not properly account for the harms caused by speech, and that it should allow for more regulation of harmful kinds of speech. This Article evaluates the various ways courts already employ free speech consequentialism. It then establishes and defends a principled basis for determining when speech's harms greatly outweigh its virtues. Courts should engage in free speech consequentialism sparingly, and should constrain themselves to considering only the harms caused by speech that can be analogized to harms caused by conduct. This Article develops a framework that recognizes the need to incorporate free speech consequentialism, and to constrain it, at various stages of First Amendment analysis, in connection with both tort and criminal law. It then applies this framework to timely and difficult speech issues, including campus hate speech, revenge porn, trigger warnings, and violent speech—with the aim of rehabilitating core values of our First Amendment doctrine and practice.

* Climenko Fellow and Lecturer on Law, Harvard Law School; Assistant Professor, Ohio Northern Law School (beginning August 2016). I would like to thank Oren Bar-Gill, Rebecca Brown, Michael Coenen, Glenn Cohen, Daniel Deacon, Daniel Epps, Richard Fallon, Daniel Farbman, George Fisher, Joseph Fishman, Charles Fried, John Goldberg, Bert Huang, Jacob Kreutzer, Benjamin Levin, Leah Litman, Brian Richardson, Susannah Barton Tobin, Mark Tushnet, and Matthew Wansley for their insights and feedback.

INTRODUCTION

Any kind of balancing test that determines whether speech may be constitutionally regulated poses serious challenges to foundational First Amendment principles. The very act of balancing tends to be both subjective and indeterminate—it is difficult to quantify the relevant harms and benefits, and it is equally difficult to sensibly weigh them against one another.[1] In the context of the First Amendment, balancing has the potential to undermine strong free speech protections and our faith in neutral principles underlying the First Amendment.[2] Courts should thus avoid determining whether speech can be regulated by balancing the harms caused by speech against the harms to speech.

---

1. See Lawrence Rosenthal, First Amendment Investigations and the Inescapable Pragmatism of the Common Law of Free Speech, 86 Ind. L.J. 1, 3 (2011) (arguing jurisprudence "that endeavors to balance the costs and benefits of a challenged regulation of speech . . . cannot be performed in any principled way and would instead amount to an invitation for unbridled judicial activism"); see also Mark V. Tushnet, Anti-Formalism in Recent Constitutional Theory, 83 Mich. L. Rev. 1502, 1509–18 (1985) [hereinafter Tushnet, Anti-Formalism] (offering in-depth critique of balancing in constitutional adjudication that explores necessity of comparing incommensurable values, arbitrariness in the way values are characterized, and difficulties in defining boundaries of side constraints on utilitarianism).

2. See Eugene Volokh, How Harassment Law Restricts Free Speech, 47 Rutgers L. Rev. 563, 574–79 (1995) (discussing how balancing gives judges greater discretion and how categorical formalism may be more speech protective).

But that is not always possible. A jurisprudence that never engaged in balancing would be absolutist in ways that both overprotect and underprotect speech. Some kind of consideration of speech harms and benefits to determine which speech is constitutionally protected—what this Article terms "free speech consequentialism"[3]—is unavoidable at some point in the First Amendment analysis.[4] Even if courts look first to purposivist or rights-based considerations,[5] courts inevitably confront their views about the value of the speech at issue in relation to the harms it causes, often in light of their theories about the instrumental goals of the First Amendment.

This Article argues that courts should constrain free speech consequentialism by considering only the speech harms that are sufficiently similar to conduct harms when evaluating the harms caused by speech. Speech harms typically have unique properties, such as being context dependent and caused by diffuse parties, but some harms caused by speech resemble the more direct and immediate harms arising in paradigmatic cases of conduct.[6] Analogizing speech harms to conduct harms would allow courts to protect individuals from the more tangible harms caused by speech while preserving the specialness of speech's virtues. After describing and justifying this constrained approach to free speech consequentialism, this Article then applies the proposal to analyze timely and difficult free speech issues.

Strong free speech protections come at the expense of many types of speech-related harms, including emotional distress, privacy intrusions, reputational damage, and violence provoked in audiences. A recent wave of scholarship argues for more explicit and more heavy-handed forms of free speech consequentialism to remedy these harms.[7] Scholars have

---

3. Free speech consequentialism "would support limitations on our right to free speech if they, for example, produced a society enjoying more overall happiness." Timothy P. Terrell, Statutory Epistemology: Mapping the Interpretation Debate, 53 Emory L.J. 523, 539 (2004). This fits within the general consequentialist approach to the law or to morality, which "[r]oughly speaking . . . means that alternative courses of action are evaluated according to the desirability of their consequences." Lea Brilmayer, Rights, Fairness, and Choice of Law, 98 Yale L.J. 1277, 1285 (1989). In order for free speech consequentialism to be a meaningful concept, this Article limits its parameters to the balancing of harms and benefits as a way of determining or applying First Amendment doctrine. The concept does not include consequentialist versus nonconsequentialist reasons for the acceptance of the First Amendment in the first place, or acceptance of the Constitution more generally.

4. See Stanley Fish, Fraught with Death: Skepticism, Progressivism, and the First Amendment, 64 U. Colo. L. Rev. 1061, 1086 (1993) (arguing because no one is truly a free speech absolutist, and "[t]here is always an exception up one's sleeve," balancing is unavoidable on some level, and "everyone is a consequentialist").

5. See infra section I.C (discussing purposivist alternative to free speech consequentialism).

6. See infra Part II (discussing manifestations of speech harms).

7. See Alan K. Chen & Justin Marceau, High Value Lies, Ugly Truths, and the First Amendment, 68 Vand. L. Rev. 1435, 1471–80 (2015) (arguing in favor of protection of "investigative deceptions," or lies told to gain access to information that exposes illegality

begun to criticize free speech jurisprudence for being dismissive of harm, and for not properly distinguishing the different mechanisms by which speech causes harm.[8] Although the First Amendment currently occupies a vaunted position in our legal and cultural practices, scholars have begun to use arguments sounding in consequentialism to chip away at the rules-based First Amendment regime.

Scholars who espouse explicitly consequentialist theories of the First Amendment believe that free speech's value lies in advancing particular ends, such as truth or democratic self-government.[9] These free speech consequentialists argue that speech can and should be suppressed when a given instance of speech actually works against those ends, or, more generally, when the benefits of that speech are outweighed by other harms.[10] These scholars advocate for a variety of approaches to balancing speech rights against other interests and may have expansive or limited conceptions of which speech should be protected, but they share the view that speech is valuable to the extent it achieves particular ends.[11] In recent years, based in part on technological advancements that facilitate speech harm, scholars have argued for the regulation of revenge porn,[12]

---

or other important information, as high value form of expression); David S. Han, The Mechanics of First Amendment Audience Analysis, 55 Wm. & Mary L. Rev. 1647, 1682–83 (2014) (arguing foreseeability of audience processing speech should be considered harm worthy of judicial remedy); Frederick Schauer, Harm(s) and the First Amendment, 2011 Sup. Ct. Rev. 81, 82 [hereinafter Schauer, Harm(s) and the First Amendment] (arguing Supreme Court has been complicit in downplaying harms caused by speech); Rebecca L. Brown, The Harm Principle and Free Speech 6 (Univ. of S. Cal. Gould Sch. of Law, Ctr. for Law & Soc. Sci. Legal Studies Research Papers Series, Working Paper No. 15-11, 2015), http://ssrn.com/abstract=2584080 (on file with the *Columbia Law Review*) (arguing in order to allow government more leeway in regulating speech harms, exacting scrutiny applied to content-based regulations should be abandoned where speech does not engage audience's rational processes).

8. See Schauer, Harm(s) and the First Amendment, supra note 7, at 83–84 (distinguishing three different ways in which speech causes harm so "we may better understand this delicate relationship between harm and free speech, and perhaps begin to understand how the doctrine might accommodate it").

9. See, e.g., Chen & Marceau, supra note 7, at 1438 (explaining explicitly consequentialist view of First Amendment supports argument that certain "high value" lies should be protected because such deceptions affirmatively further free speech goals of "enhancing political discourse, revealing truth, and promoting individual autonomy").

10. See infra section I.A (surveying consequentialist literature).

11. See Ronald J. Krotoszynski, Jr., Free Speech Paternalism and Free Speech Exceptionalism: Pervasive Distrust of Government and the Contemporary First Amendment, 76 Ohio St. L.J. 659, 665–68 (2015) (describing how Martin Redish and Cass Sunstein both argue First Amendment has instrumental value to democratic self-government, but Redish argues in favor of expansive protections for speech, while Sunstein's theory allows for fine-grained determinations of when government suppression will actually facilitate democratic self-government).

12. See, e.g., Mary Anne Franks, Sexual Harassment 2.0, 71 Md. L. Rev. 655, 684 (2012) (arguing for expansion of sexual harassment law to internet sites); Elizabeth M. Ryan, Note, Sexting: How the State Can Prevent a Moment of Indiscretion from Leading to a Lifetime of Unintended Consequences for Minors and Young Adults, 96 Iowa L. Rev. 357,

cyberbullying and internet harassment,[13] and the disclosure of true details about people's identities or locations,[14] based on the perceived minimal benefits associated with these kinds of speech as compared to the substantial harms such speech generates. Scholars often advocate for speech regulations to occur through tort law, which allows private values to be weighed against First Amendment concerns.[15]

Even scholars who favor what they deem nonconsequentialist theories of free speech, and who believe, for example, that free speech has inherent value and is a right of autonomous moral agents,[16] will in some circumstances balance these values against the harms speech causes. This balancing would occur for so-called nonconsequentialists either in defining what constitutes speech, in determining which categories of speech are protected, or in evaluating whether speech that is protected can nonetheless be prohibited because its harms greatly outweigh its virtues.[17] Some scholars would argue that free speech rights are balanced not against harms but against other rights, such as the right to privacy, property, or reputation. However, unless one of the rights at issue is defined absolutely, resolving this conflict would also require consideration of the harms at issue and the value of the speech. Thus, the question becomes not whether free speech consequentialism is appropriate, but how harms caused by speech should be accounted for in First Amendment jurisprudence.

The allure of free speech consequentialism is also reflected in the courts. Describing the Supreme Court's approach to content-based

---

381 (2010) (arguing states should recognize intentional infliction of emotional distress claims by victims of revenge porn).

13. See, e.g., Nancy S. Kim, Web Site Proprietorship and Online Harassment, 2009 Utah L. Rev. 993, 1008–12 ("While existing remedies . . . may adequately address offline harassment, they are inadequate to deal with online harassment for several reasons."); Shira Auerbach, Note, Screening Out Cyberbullies: Remedies for Victims on the Internet Playground, 30 Cardozo L. Rev. 1641, 1667–74 (2009) (suggesting defamation and intentional infliction of emotional distress as potential avenues of recovery for cyberbullying victims).

14. See Ashutosh Bhagwat, Details: Specific Facts and the First Amendment, 86 S. Cal. L. Rev. 1, 34 (2012) (arguing specific factual details are not as important to First Amendment values as ideas or more general facts, and thus should be more easily regulated).

15. See Deana Pollard Sacks, Constitutionalized Negligence, 89 Wash. U. L. Rev. 1065, 1083–85 (2012) (proposing negligence scheme for speech that unreasonably causes injury based on balancing nature of speech, vulnerability of plaintiffs, and state's interest in punishing speech).

16. See Fish, supra note 4, at 1085 (arguing speech has inherent value and is "an essential . . . feature of a just . . . society" as right of autonomous moral agents). Because even these so-called nonconsequentialist views of free speech must account for the harms speech causes, this Article will focus on when balancing should occur in a universe where free speech consequentialism is unavoidable; this is a separate inquiry from whether the justification for protecting free speech in the first place should be instrumental or nonconsequentialist.

17. See infra section II.C (addressing objections to free speech consequentialism).

restrictions on speech is superficially simple. Laws that suppress speech on the basis of content are subject to the strictest constitutional scrutiny, which is often outcome determinative.[18] Strict scrutiny is a demanding standard.[19] But in operation, the doctrine is much more complex—it incorporates considerations of harm in multiple ways. In a variety of cases, different groups of concurring and dissenting Justices have shown willingness to relax the strict scrutiny applied to content-based restrictions in order to account for the harm from depictions of animal cruelty,[20] violent video games,[21] and lies about military honors.[22] The Supreme Court is not even clear on at what point in its First Amendment analysis, or at what level of abstraction, this balancing should be performed, if at all, when free speech doctrine intersects with both criminal and tort law.[23]

---

18. See Geoffrey R. Stone, Content-Neutral Restrictions, 54 U. Chi. L. Rev. 46, 48 (1987) [hereinafter Stone, Content-Neutral Restrictions] ("Indeed, outside the realm of low-value speech, the Court has invalidated almost every content-based restriction that it has considered in the past thirty years."); see also Case Comment, First Amendment—Freedom of Speech—Content Neutrality—*McCullen v. Coakley*, 128 Harv. L. Rev. 221, 221 (2014) ("With few exceptions, the Court has struck down laws that make facial distinctions between different subject matters or viewpoints, while upholding those that do not.").

19. In recent years, the Court has struck down a wide variety of content-based restrictions. See Brown v. Entm't Merchs. Ass'n, 131 S. Ct. 2729, 2738 (2011) (striking down criminal prohibition on sale of violent video games to minors); Snyder v. Phelps, 131 S. Ct. 1207, 1219 (2011) (overturning civil damage award for vitriolic protests at military funeral); United States v. Stevens, 559 U.S. 460, 482 (2010) (striking down criminal prohibition on animal crush pornography as overbroad). But see Holder v. Humanitarian Law Project, 561 U.S. 1, 26–30, 40 (2010) (upholding law banning provision of material aid, including legal advice and support for lawful activities, to designated foreign terrorist organizations despite applying strict scrutiny).

20. See *Stevens*, 559 U.S. at 483 (Alito, J., dissenting) (rejecting application of overbreadth doctrine and leaving it to court of appeals to decide whether videos sold by respondent were constitutionally protected).

21. See *Entm't Merchs. Ass'n*, 131 S. Ct. at 2747 (Alito, J., concurring in the judgment) (arguing law attempting to prevent minors from purchasing video games should not have to satisfy strict scrutiny); id. at 2765–66 (Breyer, J., dissenting) (arguing against mechanical application of strict scrutiny and in favor of looking to "whether, overall, the statute works speech-related harm . . . out of proportion to the benefits that the statute seeks to provide" (internal quotation marks omitted) (quoting United States v. Playboy Entm't Grp., 529 U.S. 803, 841 (2000) (Breyer, J., dissenting))); id. at 2762 (Breyer, J., dissenting) ("In determining whether the statute is unconstitutional, I would apply both this Court's 'vagueness' precedents and a strict form of First Amendment scrutiny.").

22. See United States v. Alvarez, 132 S. Ct. 2537, 2551 (2012) (Breyer, J., concurring in the judgment) (rejecting "strict categorical analysis" in favor of basing "conclusion upon the fact that the statute works First Amendment harm, while the Government can achieve its legitimate objectives in less restrictive ways").

23. The Court's division in *United States v. Alvarez* over at what stage—if at all—to incorporate free speech consequentialism illustrates its stubborn problems. See Rodney A. Smolla, Categories, Tiers of Review, and the Roiling Sea of Free Speech Doctrine and Principle: A Methodological Critique of *United States v. Alvarez*, 76 Alb. L. Rev. 499, 509–19, 526 (2013) (arguing approaches in *Alvarez* all fail to articulate clear free speech principle and fate of whether free speech will be protected robustly "as in democratic elections and the actions of legislative bodies . . . all comes down to votes"). But see infra section II.A

For example, *United States v. Alvarez* recently invalidated a federal criminal law that prohibited lying about receiving a military honor.[24] The plurality opinion, authored by Justice Kennedy, disavowed free speech consequentialism at all levels of abstraction (including the highest level that determines what constitutes protected speech), proclaiming that content-based restrictions on speech are subject to strict scrutiny unless the speech fits into a small number of historically unprotected categories.[25] But while the plurality denounced free speech consequentialism in determining the value of particular speech and the resulting protections afforded to it, the Court un-self-consciously used some form of free speech consequentialism—the strict scrutiny framework the plurality adopted required it to weigh the government's interests in regulating speech against the costs of doing so, with a heavy thumb on the scale in favor of speech.[26] The concurring opinion, authored by Justice Breyer, embraced an even stronger version of free speech consequentialism. Justice Breyer's concurrence first adopted a threshold balancing test, weighing the benefits and harms of the speech in question to determine whether the criminal prohibition would be reviewed under a more relaxed form of constitutional scrutiny, which itself would then also incorporate free speech consequentialism.[27]

If even Justice Kennedy's categorical treatment of content-based restrictions requires a strict scrutiny that encourages policy-laden consequentialist balancing of the costs and benefits of imposing liability for speech, free speech consequentialism is pervasive and unavoidable. Despite the scholarly trend toward more explicit free speech consequentialism, there has been no scholarly account that proposes ways to incorporate free speech consequentialism into First Amendment doctrine in a principled way that honors strong free speech protections.

This Article assumes that strong free speech protections are an extraordinary and essential part of American First Amendment law,[28] and

---

(arguing there are principled ways of reconciling free speech consequentialism with categorical approach).

24. See *Alvarez*, 132 S. Ct. at 2551 (Kennedy, J.) (plurality opinion) (holding Stolen Valor Act infringes First Amendment protected speech).

25. See id. at 2544 (Kennedy, J.) (plurality opinion) ("[C]ontent-based restrictions on speech have been permitted, as a general matter, only when confined to the few 'historic and traditional categories [of expression] long familiar to the bar.'" (quoting *Stevens*, 559 U.S. at 468) (majority opinion)).

26. See infra notes 163–164 and accompanying text (explaining how free speech consequentialism can be incorporated into free speech doctrine at various levels of abstraction).

27. See *Alvarez*, 132 S. Ct. at 2553 (Breyer, J., concurring in the judgment) (noting how, although "false factual statements enjoy little First Amendment protection," they also "serve useful human objectives" in both social and public contexts).

28. See Timothy Zick, First Amendment Cosmopolitanism, Skepticism, and Democracy, 76 Ohio St. L.J. 705, 712 (2015) (noting America's "exceptional First Amendment free speech protections").

that strong protections for liberty are an essential part of the U.S. constitutional system generally.[29] America has uniquely expansive free speech protections, even for the most intolerant, offensive speech.[30] The scholarly trend toward free speech consequentialism, working in tandem with our society's increased attention to the emotional harms caused by speech,[31] may create a constitutional moment that threatens America's cultural commitment to protecting speech from government intrusion. However, even those who believe our current First Amendment doctrine is overprotective of speech must find ways to add coherence and clarity to free speech consequentialism.

Largely focusing on content-based restrictions on speech affecting the general public,[32] this Article argues that consequentialist accounts of harm balancing require a proper understanding of how speech harms differ from other types of harms. The harms particularly associated with speech are fundamentally different from harms caused by conduct. In order to constrain free speech consequentialism, courts should allow only those certain types of speech harms that more closely mirror conduct harms to be subject to any consequentialist balancing.

This approach is justified by both the preference for speech articulated in the First Amendment and the uniqueness of speech's properties. The First Amendment's preference for speech eschews strong forms of consequentialism and marks speech as special in both its benefits and the harms it causes. Further, because the harms caused by speech in particular are generally context dependent, bound up in our identities, malleable over time, and perpetrated by diffuse parties in addition to the speaker,[33] there are unique difficulties presented by efforts to regulate speech. Identifying how speech causes harm is more complex, and the harms associated with speech are also inextricably intertwined with

---

29. James MacGregor Burns et al., Government by the People 12 (15th ed. 1993) (describing American Constitution as being predicated on distrust for government and desire to protect individual liberties).

30. See infra notes 274–278 and accompanying text (describing America's unique protections for what other countries regulate as "hate speech").

31. See Erica Goldberg, Emotional Duties, 47 Conn. L. Rev. 809, 811 (2015) (describing recent developments that place greater emphasis on distinction between physical and emotional injuries).

32. See Stone, Content-Neutral Restrictions, supra note 18, at 47 ("Content-based restrictions limit communication because of the message it conveys. Laws that prohibit seditious libel, ban the publication of confidential information, forbid the hiring of teachers who advocate violent overthrow of the government, or outlaw the display of the swastika in certain neighborhoods are examples of content-based restrictions."); see also Geoffrey R. Stone, Content Regulation and the First Amendment, 25 Wm. & Mary L. Rev. 189, 189–90 (1983) (explaining how content-neutral restrictions target speech for elements unrelated to content, such as volume of speech or location in which it is uttered).

33. But see Brian Leiter, Cleaning Cyber-Cesspools: Google and Free Speech, *in* The Offensive Internet 156 (Saul Levmore & Martha C. Nussbaum eds., 2010) (arguing heightened harm results from Internet speech because it is permanent, divorced from context, and available to anyone).

speech's virtues: the ability to allow us to reconsider what actually constitutes a harm, to change our perception of harms, and to convince others to change their notions of harm. As a result, instead of subjecting all speech to consequentialist balancing, courts should allow a weighing of costs and benefits only of the harms more similar to conduct than classic speech. Because of our constitutional and cultural commitment to free speech, speech should be defined in a way that preserves its specialness, and some harms caused by speech simply cannot be remedied.

This Article is not just about rules versus standards, or absolutism versus balancing. It is about when and which types of harms should be accounted for in First Amendment doctrine—and about when courts may determine that the harmful consequences of speech outweigh its benefits. The term free speech consequentialism signifies that, regardless of the nature of First Amendment rights, judges must confront the proper treatment of harms in our First Amendment calculus.

This Article proceeds in three parts. Part I surveys the scholarly calls for increased incorporation of free speech consequentialism and demonstrates how pervasive and inevitable free speech consequentialism already is within First Amendment doctrine. Part I also compares the different ways to incorporate free speech consequentialism. Next, Part II proposes ways to discipline and constrain free speech consequentialism. The types of speech harms more similar to conduct are more amenable to a balancing of harms and benefits than other, more intangible and emotional, harms caused by speech. Finally, Part III applies this framework to examples such as campus hate speech, revenge porn, trigger warnings, and violent or dangerous speech.

I. THE INEVITABLE MOVE TOWARD FREE SPEECH CONSEQUENTIALISM

This Part introduces the idea of free speech consequentialism. Section I.A defines the idea and surveys the recent scholarly interest in favor of free speech consequentialism. Section I.B then shows that free speech consequentialism is in fact more pervasive than many scholars realize. Free speech consequentialism appears in various First Amendment doctrines ranging from the existence of categories of unprotected speech to the prior restraint doctrine. Section I.C argues that free speech consequentialism, more than being ubiquitous, is in fact inevitable. Finally, section I.D examines the different ways judges approach free speech consequentialism and discusses the virtues and vices of these approaches.

A. *Calls for Free Speech Consequentialism*

First Amendment scholars have increasingly called for First Amendment doctrine to move toward different forms of free speech consequentialism.[34] These scholars believe that courts should, in various

34. See supra notes 7–17 (surveying consequentialism scholarship).

ways, balance speech's harms against its benefits, often in light of some instrumental purpose of the First Amendment, in order to determine whether a variety of speech-restrictive laws are unconstitutional.[35] Free speech consequentialist scholars often take as a given that the scope of the First Amendment should be determined by whether the harms of speech are outweighed by its virtues, but differ on how courts should weigh speech's harms and virtues and on which consequences are so dire that they necessitate the regulation of speech.

To advance more sophisticated forms of free speech consequentialism, one group of scholars has begun to establish frameworks for conceptualizing and differentiating the various harms caused by speech. These scholars are motivated by the view that the Supreme Court treats the harms caused by speech dismissively without genuinely assessing the evidence.[36] Frederick Schauer's "trilogy of harms," for example, separates the harms caused by speech into (1) harms of advocacy, where speech influences a third party to commit a harmful act;[37] (2) harms of "verbal assault," where the words themselves have a negative effect on those who hear or see the speech;[38] and (3) participant harms, where the primary harm flows to a participant (who may be either willing or unwilling) in the production of the speech.[39] To Schauer's trilogy of harms, David Han adds "dissemination-based harms," or harms that flow "from the mere fact that the information was disseminated to the listener, often with the general assumption that the information disclosed will continue to be passed on to others."[40] Defamation and libel fall into this category, along with speech that is constitutionally protected but might harm others' desire for privacy or control over their reputations.

Disaggregating these types of harms is necessary, according to Schauer, to "enable courts to determine the extent of the harms involved,

---

35. See Alexander Tsesis, Balancing Free Speech, 96 B.U. L. Rev. (forthcoming 2016) (manuscript at 3–4) (on file with the *Columbia Law Review*) (contending balancing approach best satisfies purpose of First Amendment, which supports individual right to free speech in order to serve common good).

36. Han, supra note 7, at 1657 ("[T]he Court has given relatively short shrift to the issue of speech-based harm . . . ." (emphasis omitted)); Schauer, Harm(s) and the First Amendment, supra note 7, at 81–83 ("[The] Supreme Court has often been complicit in denying or downplaying the harm-producing capacity of speech . . . .").

37. See Schauer, Harm(s) and the First Amendment, supra note 7, at 97–100 (using case involving dangers to children of violent games, Brown v. Entm't Merchs. Ass'n, 131 S. Ct. 2729 (2011), to illustrate how speech influences third party to commit harmful act).

38. See id. at 100–02 (invoking case involving Westboro Baptist Church's inflammatory protest of military funeral, Snyder v. Phelps, 131 S. Ct. 1207 (2011), to illustrate direct emotional harm of speech); see also *Snyder*, 131 S. Ct at 1220–21 (Breyer, J., concurring) (noting how "speech, like an assault, seriously harms a private individual").

39. See Schauer, Harm(s) and the First Amendment, supra note 7, at 103–04 (discussing case involving pornography produced while torturing animals, United States v. Stevens, 559 U.S. 460 (2010), to demonstrate participant harms).

40. Han, supra note 7, at 1650.

and whether the doctrine should allow any redress against them."[41] Han, building on Schauer's work, goes further and argues in favor of new harms-based First Amendment doctrine. Han believes courts should attune themselves more to audience reaction, because audience response is often the mechanism through which speech causes harm.[42] He proposes that the Supreme Court use empirical evidence and a predictive approach to assess how an average listener would actually react to speech, including whether an average listener would actually perceive speech as threatening or hurtful.[43] When crafting doctrine that categorizes unprotected speech as a true threat or as incitement, courts "should measure the social harm resulting from the audience's processing of particular speech based on what the speaker should have reasonably foreseen under the circumstances."[44] Han's empirical, consequentialist way of incorporating speech harms contrasts with what he believes to be an unsubstantiated normative approach, advocated by scholars like Larissa Lydsky, in which audiences are assumed to be rational, skeptical, thick-skinned, and capable of sorting out lies from the truth.[45]

Another group of scholars argues that the Court should embrace free speech consequentialism by adopting a more permissive approach toward allowing tort law to regulate speech. These scholars claim that tort law is more amenable to balancing the costs and benefits of speech, because criminal law is more coercive and less finely tailored.[46] According to Deana Pollard Sacks, "[t]he Court has constitutionalized a variety of tort claims arising from speech and created a balancing framework for reconciling First Amendment values with the state's interest in punishing and deterring injurious speech, as opposed to applying strict scrutiny for criminal regulation of speech."[47] Sacks extols the ability of tort law to better account for the harms caused by speech and argues that the Supreme Court's invalidation, in *Brown v. Entertainment Merchants*

---

41. Schauer, Harm(s) and the First Amendment, supra note 7, at 107.

42. Han, supra note 7, at 1662.

43. See id. at 1679–81 (arguing courts "could seek to predict, in a more empirical and contextual manner, how the particular targeted audience in question would *likely* or *foreseeably* process the speech").

44. Id. at 1697.

45. See id. at 1684–85 (noting Lydsky believes courts should adhere to "rational audience" assumption, which presumes audiences are "non-impulsive, thick-skinned, and generally resistant to processing speech in harmful ways"). Although Han agrees with Lydsky that these normative ideals are important in justifying the philosophical underpinnings of First Amendment analysis, he disagrees that they should be used to assess the harms caused by speech. See id. at 1685–86 ("But the mere fact that the broad philosophical underpinnings of much of our First Amendment jurisprudence assume a rational audience does not necessarily mean that courts should calculate the harm resulting from speech based on an idealized rational audience construct.").

46. See, e.g., Sacks, supra note 15, at 1067–68 ("Tort liability is inherently more narrowly tailored than criminal regulation . . . .").

47. Id. at 1067.

Association,[48] of a criminal ban on the sale of violent video games to children "does not control the issue of negligence liability for actual harm caused by children's consumption of violent video games."[49] Criminal and tort law should differ on this issue, according to Sacks, because criminal penalties require no proof of harm.[50] By contrast, tort law can tailor itself to combat actual harm,[51] using scientific evidence of how children's brains are impacted by violent video games.[52] (It is unclear whether this would/should count as an actual injury as far as tort law is concerned, but Sacks's point stands regardless.)

Sacks argues that criminal penalties for selling violent video games, unlike tort penalties based on unreasonably dangerous speech, "risk[] chilling speech that does not actually cause harm."[53] Because Sacks believes the *Brandenburg* incitement test—which allows speech to be regulated if it is directed to and likely to cause imminent violent action[54]—is incapable of balancing the benefits and costs of speech that causes real harm,[55] she proposes to allow negligence liability for unreasonably dangerous speech.[56] Her proposal would condition liability upon "public policy analyses concerning the media's propensity to influence the audience, the likelihood and gravity of public risk, and whether the risks are sufficiently foreseeable to warrant punishment."[57]

While Sacks locates negligence as an area for increased tort liability for harmful speech, Nathan Oman and Jason Solomon seek to use tort law to bolster intentional infliction of emotional distress (IIED) claims for speech-based harms.[58] Oman and Solomon take issue with the Supreme

---

48. 131 S. Ct. 2729, 2741–42 (2011).

49. Sacks, supra note 15, at 1068.

50. See id. ("Criminal penalties for sales require no proof of actual harm to be enforced and risks chilling speech that does not actually cause harm.").

51. See id. ("Tort liability is inherently more narrowly tailored than criminal regulation by virtue of the proof elements necessary to establish a claim in a particular case, including proof of actual harm to the plaintiff and both factual and proximate causation.").

52. See id. at 1072, 1074–76 ("Violent video games are distinguishable from other forms of violent media such as television, motion pictures, and music because they are interactive and repetitive.").

53. Id. at 1068.

54. See Brandenburg v. Ohio, 395 U.S. 444, 447 (1969) ("[C]onstitutional guarantees of free speech and free press do not permit a State to forbid or proscribe advocacy of the use of force or of law violation except where such advocacy is directed to inciting or producing imminent lawless action and is likely to incite or produce such action.").

55. See Sacks, supra note 15, at 1133 (noting immunizing certain speech from regulation "shifts the costs of harm from the speech producers to the public at large and may increase the sum total of social harm").

56. Id. at 1068.

57. Id. at 1118.

58. See Nathan B. Oman & Jason M. Solomon, The Supreme Court's Theory of Private Law, 62 Duke L.J. 1109, 1135–36 (2013) ("The intentional infliction of emotional

Court's decision to equate criminal law and tort law when assessing First Amendment protections.[59] They argue that tort law, as private law,[60] is better conceptualized as a mechanism for individual justice than as a mechanism for state action.[61] Courts should balance plaintiffs' rights to individual redress against free speech values by "containing the right to civil recourse rather than cutting it off altogether."[62] According to Oman and Solomon, the Supreme Court erred in its speech-torts jurisprudence by treating tort law, where individuals seek redress for private wrongs, as a subspecies of public regulation, or just "a tool used by government to suppress and punish speech."[63]

Thinking of tort law as governmental regulation led the Court in *Snyder v. Phelps* to invalidate a claim for IIED brought by a grieving father against the offensive protesting of the Westboro Baptist Church near his son's funeral.[64] Oman and Solomon take issue with the fact that "Chief Justice Roberts's majority opinion in *Snyder* does not even attempt to articulate a justification for state tort law, instead focusing the bulk of its discussion on the nature of Westboro's speech."[65] Oman and Solomon argue in favor of balancing plaintiffs' need for redress with First Amendment interests in a way that is less binary than a liability or no-liability regime.[66]

Finally, many scholars take it for granted that discussions of what speech should be protected comes down to questions of harms, or costs and benefits. Free speech consequentialism is justified by the view that courts should not simply ignore the harms caused by speech, especially in cases of low-value speech or speech that does not engage one rationally. For example, Rebecca Brown argues that the Court's current strict-scrutiny framework disables the government from addressing the real

---

distress tort provides a clear example of a tort that was created by judges to provide redress for victims of wrongs, and in doing so, to reinforce social equality.").

59. Id. at 1111 (arguing "Supreme Court itself failed to appreciate the private-law nature of" Snyder v. Phelps, 131 S. Ct. 1207 (2011), where Court invalidated damages award for IIED on First Amendment grounds).

60. By "private law," Oman and Solomon mean "common-law subjects like torts, contract, and property (and their statutory counterparts) that involve primary rights by individuals that can be enforced by the rights-holders themselves against other individuals and entities." Id. at 1111–12.

61. See id. at 1119 ("In opposition to the instrumentalist paradigm of private law, an alternative view . . . holds that private law is about individual justice.").

62. Id. at 1114.

63. Id. at 1109, 1112–13.

64. See id. at 1137 (attributing Court's decision to assumption tort law is "species of government regulation").

65. Id. at 1134.

66. Id. at 1161–67.

harms caused by speech.[67] Her free speech consequentialism involves a "harm principle" that gives the government more leeway in regulating speech if the government does not have a "censorial" motive, but is instead regulating speech because of the harm it produces.[68]

Brown proffers a principled theory that speech may be regulated when it causes harm unrelated to persuasion through rational thinking processes.[69] However, Brown's theory would allow the government to avoid the strict scrutiny currently applied to content-based restrictions on speech so long as the government can demonstrate a theory of harm that is unrelated to stifling ideas, such as protecting privacy, reducing violence, or safeguarding vulnerable individuals from social ills.[70] But many, if not most, attempts to censor speech can be converted into, or may have as a secondary or even primary goal, an articulation that the regulation is limiting the harm caused by speech.[71] Ultimately, her approach would likely erode the strong protections of our current free speech framework based on the government's ability to prove that it is pursuing equality or alleviating the nonexpressive harms caused by speech instead of targeting the harms related to the message of the speech.[72]

Other scholars looking to cabin free speech absolutism offer new exceptions to protection for speech they find particularly harmful. Those with broad concepts of free speech consequentialism believe that courts

---

67. Brown, supra note 7, at 6–7 (arguing near absolute protection for speech like entertainment, aids to sexual gratification, or "hateful harangues against individuals or groups" are "backwards").

68. See id. at 10–12 ("Laws shown to act non-censorially to address a real harm should receive the lesser scrutiny afforded to content-neutral laws.").

69. See id. at 34 ("Non-censorial restrictions, on the other hand, allege that harm will result from expression by virtue of some kind of non-rational impact that the expression can be expected to have by its utterance, not dependent on any idea or message that it may also contain.").

70. See id. at 12 (arguing "government should have a meaningful opportunity to demonstrate to the Court that it is reasonably protecting against social harms validly within the police power").

71. See id. at 57 ("If the government sought to prohibit speech because it was offensive and caused distress to individuals, its restriction would rest on a censorial theory of harm and would have to be subjected to strict scrutiny. This is tantamount to the proscription of bad or offensive ideas."); see also id. at 57–58 (noting government could offer noncensorial motive if it sought "to supply a public culture in which all persons are assured of their status as citizens in good standing").

72. Brown may be correct that her proposal will not change current doctrine much. See id. at 13–14 (noting change in jurisprudence under proposal may not be "vast"). But allowing the government to avoid strict scrutiny in cases where the government can articulate a theory of harm, unrelated to a "censorial motive," could potentially undermine the exceptional free speech protections characteristic of America's First Amendment jurisprudence, because of the shift in how courts conceptually view the First Amendment. Brown favors a new conception of "ordered liberty" that is not liberty against the government but liberty from one citizen against another that gives the government far more power to regulate harm. See id. at 62–64 ("Ordered liberty describes the balance that the nation has struck between the rights of the individual and the need for order.").

should judge regulations and doctrine by their ability to maximize social welfare.[73] Narrower notions of free speech consequentialism also call for the regulation of speech if protecting the speech actually undermines a particular instrumental aim of the First Amendment, for example if protecting lies impedes the search for truth.[74] Both versions of free speech consequentialism are problematic because each individual has her own sense about which speech would be particularly or unfairly harmful, which speech is particularly low value, and which speech actually undermines the goals of the First Amendment.[75]

Within only the last few years, there have been calls for greater regulation of speech that causes harm in its provision of both true factual details[76] and lies,[77] with scholars arguing that these forms of speech are not beneficial. Brian Leiter, who argues in favor of a strong version of free speech consequentialism, believes that "most non-mundane speech people engage in is largely worthless, and the world would be better off were it not expressed."[78] The Internet, a medium that provides unprecedented access to speech for those without great resources, is the locus of calls for reforms that suppress speech,[79] especially anonymous speech, by some feminists, who believe that misogynist reactions to blogs place a "tax" on women seeking to participate in Internet conversations.[80]

---

73. See Lawrence B. Solum, Public Legal Reason, 92 Va. L. Rev. 1449, 1501 (2006) (arguing "public legal reason" should buttress normative legal theory and supports "both welfare and fairness").

74. Helen Norton, Lies and the Constitution, 2012 Sup. Ct. Rev. 161, 163, 172–78 (arguing presumption that lies ought to be fully protected "should not govern a specific category of low-value lies that themselves undermine First Amendment interests"); see also Fish, supra note 4, at 1085 ("[I]f you have . . . a consequentialist view of the First Amendment—a view that values free speech because of the good effects it will bring about—then you must necessarily be on the lookout for forms of action, including speech action, that threaten to subvert those effects.").

75. See Ronald Turner, Regulating Hate Speech and the First Amendment: The Attractions of, and Objections to, an Explicit Harms-Based Analysis, 29 Ind. L. Rev. 257, 292 (1995) (discussing how majority and dissenting opinions in Texas v. Johnson, 491 U.S. 397 (1989), differed as to whether flag burning has tendency to incite breach of peace).

76. Bhagwat, supra note 14, at 61 (arguing specific factual details are not as important to First Amendment values as ideas or more general facts, and thus should be more easily regulated).

77. See Norton, supra note 74, at 163 (arguing certain types of lies can be prohibited by government "consistent with the First Amendment").

78. Brian Leiter, The Case Against Free Speech 3 (Univ. of Chi. Pub. Law & Legal Theory, Working Paper No. 481, 2014) [hereinafter Leiter, Case Against Free Speech], http://ssrn.com/abstract=2450866 (on file with the *Columbia Law Review*). Leiter defines nonmundane speech as speech that is separate from the mundane details that allow us to plan our daily lives. Id. at 3.

79. See Franks, supra note 12, at 699 (arguing sexual harassment law should be extended to cyberspace).

80. See, e.g., Cheryl Abbate, Gender-Based Violence, Responsibility, and John McAdams (Jan. 20, 2015), http://ceabbate.wordpress.com/2015/01/20/gender-based-violence-responsibility-and-john-mcadams/ [http://perma.cc/N7GS-7QG8] (discussing

Many scholars who advance these consequentialist arguments focus not only on maximizing welfare, but on distributional concerns. Schauer notes the distributional concern that the costs of speech are "rarely borne equally, or even fairly" throughout the population.[81] Scholars like Danielle Citron and Mary Anne Franks argue that harms from speech that disproportionately befall women are more likely to be minimized,[82] compounding the problem of distributional inequity of speech harms. Concerns with the distribution of harms or concepts like fairness could be considered deontological, not consequentialist, values.[83] However, consequentialism as a moral philosophy focuses on distribution as part of overall outcome, with the needs of the disadvantaged sometimes adding more to total welfare due to diminishing marginal returns on welfare.[84] Plus, most scholars concerned with fairness still resort to weighing speech harms to particularly vulnerable populations against their benefits when arguing in favor of greater regulation of speech.

If courts followed the lead of scholars and began to seriously evaluate empirical evidence of the harms of speech and devalue its benefits, America's exceptional commitment to strong free speech protections would be greatly undermined.[85] There is a greater chance for courts, based on their own subjective views or ideological priors, to decide that certain speech, even core speech, is too harmful to be tolerated. Perhaps the only principled way to deal with speech harms is to minimize them, assume that individuals can largely manage them (in contrast to harms caused by conduct), and believe that the marketplace of ideas remedies them.

B. *Free Speech Consequentialism as Doctrine*

Scholars are correct that the Supreme Court is dismissive of speech harms. The Court does not have a sophisticated framework for dis-

---

gender implications of anonymous postings). Of course, tax is not exactly the right framing, given that private individuals' speech, and not the government, is the source of the burden on women who must deal with the emotional costs of hostile cyber-speech.

81. Schauer, Harm(s) and the First Amendment, supra note 7, at 110.

82. See Danielle Keats Citron & Mary Anne Franks, Criminalizing Revenge Porn, 49 Wake Forest L. Rev. 345, 347–48 (2014) ("As revenge porn affects women and girls far more frequently than men and boys, and creates far more serious consequences for them, the eagerness to minimize its harm is sadly predictable.").

83. See Solum, supra note 73, at 1496–98 (describing deontological theories).

84. Plus, some scholars argue that fairness can also be considered a consequentialist value. See Ward Farnsworth, The Taste for Fairness, 102 Colum. L. Rev. 1992, 2003 (2002) (reviewing Louis Kaplow & Steven Shavell, Fairness Versus Welfare (2002)) (arguing because people have second-order preferences for fairness, improving fairness will also improve utility). At this level of abstraction (and at higher levels of abstraction), the concept of consequentialism ceases to have useful meaning.

85. See Guy E. Carmi, Dignity—The Enemy from Within: A Theoretical and Comparative Analysis of Human Dignity as a Free Speech Justification, 9 U. Pa. J. Const. L. 957, 988–89 (2007) (describing how America differs from other Western democracies in its commitment to "robust free speech protection").

tinguishing speech harms, evaluating empirical evidence, or predicting audience reaction. That being said, a surprising number of First Amendment doctrines go unnoticed as *already* incorporating some form of free speech consequentialism. Harms and benefits are weighed when categorizing speech as high or low value, when determining whether speech fits into a particular category, and when applying the scrutiny that corresponds to particular categories of speech. This section describes those doctrines.

1. *Categories of Unprotected Speech.* — Generally, when faced with a content-based regulation, the Court first determines whether the speech at issue is "high value" or of a lesser value.[86] This categorical approach to the First Amendment itself entails forms of free speech consequentialism.

As mentioned earlier, the categorical approach subjects speech regulations to exacting scrutiny unless they are categorically exempted from protection.[87] If categorizing speech as highly protected or exempted is accomplished by weighing the costs and benefits of the speech, the exception is produced by what is referred to as "definitional balancing"[88] and implicates free speech consequentialism. If categorical exemptions are based on historical exclusions, however, or based on notions about the purpose of the First Amendment and which types of speech it was designed to protect, the Court believes it is not engaging in free speech consequentialism.[89] However, it is difficult to escape free speech consequentialism, even using a categorical approach that avoids creating new categories of unprotected speech. Even the test for obscenity, subject to a categorical exception, considers whether the speech at issue depicts sexual conduct in a "patently offensive way" and "lacks serious literary, artistic, political, or scientific value."[90]

In assessing how courts incorporate harm into First Amendment jurisprudence, it is important to recognize at the outset that a weighing of harms versus benefits occurs at various levels of abstraction,[91] which

---

86. Stone, Content-Neutral Restrictions, supra note 18, at 47–48 (noting if restricted speech does not "occupy a 'subordinate position in the scale of first amendment values,' it accords the speech virtually absolute protection" (footnote omitted) (Ohralik v. Ohio State Bar Ass'n, 436 U.S. 447, 456 (1978))).

87. See supra notes 18–19 and accompanying text (noting content-based restrictions are subject to strict scrutiny).

88. See Schauer, Harm(s) and the First Amendment, supra note 7, at 85–86 & n.27 (interpreting First Amendment coverage as being determined through "weighing the relative social costs and benefits of a particular category of speech").

89. For example, Chief Justice Roberts in *United States v. Stevens* derided as "startling and dangerous," the government's view that the creation of new categories of unprotected speech, such as speech depicting animal torture, should be based on "'a categorical balancing of the value of the speech against its societal costs.'" 559 U.S. 460, 470 (2010) (quoting Brief for the United States at 8, Stevens, 559 U.S. 460 (No. 08-769), 2009 WL 1615365, at *8).

90. Miller v. California, 413 U.S. 15, 24 (1973).

91. Judge Learned Hand used the phrase "series of abstractions" to refer to the way in which expression in a copyrighted work can be continually abstracted until it represents a

correspond to the different stages of First Amendment jurisprudence, and also at different levels of generality or particularity of the speech at issue.[92] At the earliest stage, the "coverage" stage, courts determine whether the First Amendment applies at all.[93] As examples, "[s]ecurities violations, antitrust violations, criminal solicitation, and many other categories of 'speech' remain uncovered by the First Amendment."[94] Then, at the "categorization" stage, once the Court has determined that First Amendment scrutiny applies, it decides whether the speech at issue fits into a category that is subject to a categorical exception, like obscenity or fighting words. Finally, at the last stage and usually at a lower level of generality of the speech at issue, a weighing of harms can also be performed at what can be called the "violation" stage, where the Court determines whether a particular regulation of speech that is generally protected by the First Amendment is nonetheless constitutional.

Because free speech consequentialism pervades First Amendment doctrine, the line between the categorization stage and the violation stage is blurry.[95] This is partially why even a categorical approach to First Amendment doctrine involves free speech consequentialism. Courts use free speech consequentialism both to determine which categories are unprotected, and to define those categories. For example, incitement may be regulated as unprotected speech, but speech is defined as incitement only when it is "directed to inciting or producing imminent lawless action and is likely to incite or produce such action."[96] Libelous speech is subject to a categorical exclusion from free speech protection, but different First Amendment standards determine when speech is considered libelous depending on whether the plaintiff is a public versus private figure and whether or not the speech is a matter of public concern.[97] The more

---

mere idea, not subject to copyrighted protection. Nichols v. Universal Pictures, 45 F.2d 119, 121 (2d Cir. 1930). This sense of speech being abstracted away from the particular speech at issue to whether "expression" constitutes speech at all is loosely analogous to this formulation.

92. The level of abstraction/stage of the inquiry often corresponds to the level of generality at which the speech is assessed.

93. See Frederick Schauer, The Boundaries of the First Amendment: A Preliminary Exploration of Constitutional Salience, 117 Harv. L. Rev. 1765, 1769 (2004) (distinguishing "coverage" from "protection" in First Amendment analysis by noting acts, behaviors, and restrictions that "are simply not *covered* by the First Amendment").

94. Id. at 1771.

95. See Joseph Blocher, Categoricalism and Balancing in First and Second Amendment Analysis, 84 N.Y.U. L. Rev. 375, 390 (2009) ("Many of these levels of protection [of speech] are akin to balancing, and thus it is at this stage that categorization and balancing most commonly interact: The division between types of speech is categorical, but the levels of protection are often variants of balancing." (footnote omitted)).

96. Brandenburg v. Ohio, 395 U.S. 444, 447 (1969).

97. See Michael J. Mannheimer, Equal Protection Principles and the Establishment Clause: Equal Participation in the Community as the Central Link, 69 Temp. L. Rev. 95, 141–42 & n.294 (1996) (demonstrating public-figure plaintiffs must show greater

important (and thus beneficial) the speech is to the public, the harder it is to win a lawsuit for libel based on that speech. Thus, harms balancing may have originally determined, at a high level of abstraction, that libel is a categorical exception to the First Amendment, but then harms balancing occurs again to determine which speech is actionable as libel.

2. *Tailoring and Tiers of Scrutiny.* — At the violation stage, the doctrinal structure of many First Amendment tests also reflects free speech consequentialism. Even if courts use a categorical approach that bases their determinations of which speech is unprotected on history or the original purpose of the Constitution, the constitutional scrutiny that applies to speech at various categorical levels of protection necessarily entails free speech consequentialism. Moreover, under less categorical approaches, free speech consequentialism happens both when determining the value of the speech and when applying the test to determine whether the regulation at issue survives constitutional scrutiny.

As a general matter, when faced with a content-based restriction on speech, the Court subjects regulations of high-value speech to strict scrutiny, requiring a compelling governmental interest and narrow tailoring between the goal of the statute and the means by which the law accomplishes that goal.[98] If the speech is of a lower value, the Court defines standards by which the speech may be restricted.[99] Commercial speech, which is valued somewhere between unprotected obscenity[100] and high-value speech, can be regulated if it is false or misleading.[101] Restrictions on truthful commercial speech are subject to intermediate scrutiny, and can be regulated if the restriction directly advances a substantial governmental interest and is substantially related to achieving the interest.[102] These tests assess whether there is an appropriate fit between the ends the government wishes to achieve and the means by which the government achieves those ends.

---

culpability and that presumed and punitive damages are available when speech is not matter of public concern).

98. See Stone, Content-Neutral Restrictions, supra note 18, at 48 & n.6 (explaining further, if regulation targets speech occurring on government property or limits subject matter for designed purpose, strict scrutiny is not applied).

99. See id. at 48 ("Moreover, while the Court strictly scrutinizes virtually all content-based restrictions of high-value speech, it applies a broad range of standards to test the constitutionality of content-neutral restrictions.").

100. See Miller v. California, 413 U.S. 15, 23–24 (1973) (stating obscene material, defined as appealing to "prurient interest" and lacking "serious literary, artistic, political, or scientific value," is "unprotected by the First Amendment").

101. See Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n, 447 U.S. 557, 563–64 (1980) ("The government may ban forms of communication more likely to deceive the public than to inform it, or commercial speech related to illegal activity." (citations omitted)).

102. See id. at 564 (circumscribing government regulation of commercial speech to what "directly advance[s] the state interest involved" and cannot be "served as well by a more limited restriction").

The constitutional scrutiny applied to speech is a form of free speech consequentialism because it examines both the importance of the government interest and whether the statute is tailored appropriately to serve that interest. Determining whether an interest is compelling versus significant requires courts to account for the gravity of the harms caused by speech. Then, these harms *of* speech (in the form of the government interest) are analyzed in terms of the harms *to* speech/benefits of speech (in the form of asking whether the regulation is more restrictive than necessary to accomplish its goals). This is why even means–ends tests like Justice Kennedy's approach in *United States v. Alvarez*, which explicitly rejected Justice Breyer's incorporation of free speech consequentialism, import free speech consequentialism.[103]

In *Alvarez,* which invalidated the Stolen Valor Act of 2005 and its attachment of criminal penalties to lying about the receipt of a military honor,[104] the plurality applied strict scrutiny to determine the statute's constitutionality.[105] The Court recited the compelling interests "related to the integrity of the military honors system,"[106] but concluded there was not a "direct causal link between the restriction imposed and the injury to be prevented."[107] The Court proffered that, in a free society, the way to remedy false speech is to challenge the speech.[108]

The *Alvarez* plurality, perhaps unwittingly, also marshaled free speech consequentialism when it distinguished general falsehoods, which are categorically protected speech, from perjury statutes and statutes that "prohibit falsely representing that one is speaking on behalf of the Government, or that prohibit impersonating a Government officer."[109] The plurality had to contend with the fact that certain false statements are unprotected because of the harm they cause, like harm to the integrity of government processes.[110] The Court allowed for the possibility that the falsity of speech could bear on whether it was protected while "reject[ing] the notion that false speech should be in a general category that is presumptively unprotected."[111] On some level, the Court had to compare the value and costs of the speech to determine that some false speech may

---

103. See 132 S. Ct. 2537, 2544 (2012) (plurality opinion) (Kennedy, J.) (stating balancing of relative social costs and benefits to determine level of constitutional scrutiny applied has been rejected by Supreme Court).

104. See id. at 2542 (plurality opinion) (Kennedy, J.) (explaining defendant lied at public board meeting as board member of governmental water district board about receiving Congressional Medal of Honor).

105. See id. at 2543–45 (plurality opinion) (Kennedy, J.) ("When content-based speech regulation is in question . . . exacting scrutiny is required.").

106. Id. at 2548 (plurality opinion) (Kennedy, J.).

107. Id. at 2549 (plurality opinion) (Kennedy, J.).

108. Id. (plurality opinion) (Kennedy, J.).

109. Id. at 2546.

110. See id. (plurality opinion) (Kennedy, J.) (explaining restriction on some false statements does not exempt all such restrictions from First Amendment scrutiny).

111. Id. at 2546–47 (plurality opinion) (Kennedy, J.).

be punished in the service of preventing its harms. After *Alvarez*, false speech may not be a category of unprotected speech, but the lower value of false speech (or its greater propensity to cause harm) may influence whether it can be regulated. Even a categorical approach to free speech is not as categorical as anticonsequentialist Justices might like to think.

Further, as noted above, at least two sitting Justices would apply intermediate scrutiny to deceptive speech. Justices Breyer and Kagan would assess the costs and benefits of the speech both at the categorization stage, in determining what level of constitutional scrutiny to apply, and the violation stage, where the government's ends and the means to achieve those ends are assessed.[112] The government has argued in several First Amendment cases that the Court "could create new categories of unprotected speech by applying a 'simple balancing test' that weighs the value of a particular category of speech against its social costs and then punishes that category of speech if it fails the test."[113] Accepting this approach as a way of determining the level of constitutional scrutiny afforded speech, Justice Breyer, concurring in the judgment in *Alvarez*, concluded that false speech's potential for harm and lesser value subjected the Stolen Valor Act to the more deferential intermediate scrutiny, which assesses "whether the statute works speech-related harm that is out of proportion to its justifications."[114] This type of intermediate scrutiny accounts for the harms and virtues of speech at a low level of abstraction, essentially asking whether the harms to speech are outweighed by the harms from speech.

3. *Speech Regulations that Are Entitled to More Deference.* — Although this Article focuses mainly on generally applicable, content-based restrictions on speech, it is worth noting that speech regulations that are content-neutral or target only specific subsets of the population are scrutinized by the Court using even stronger versions of free speech consequentialism. The government's ability to restrict speech on its own property, for example, depends on factors that include the cost to the government of requiring access to the speaker, the government's noncost justifications, and the degree to which the property will enhance the value of the speech.[115] The test for the constitutionality of a content-neutral restriction that targets only the time, place, or manner of an utterance

---

112. See id. at 2551 (Breyer, J., concurring in the judgment) ("In determining whether a statute violates the First Amendment, this Court has often found it appropriate to examine the fit between statutory ends and means.").

113. Brown v. Entm't Merchs. Ass'n, 131 S. Ct. 2729, 2733–35 (2011) (citing United States v. Stevens, 559 U.S. 460, 470 (2010)) (citing cases where government asked for new categories of unprotected speech to regulate violent video games, animal crush pornography, and violent speech).

114. *Alvarez*, 132 S. Ct. at 2551–52 (Breyer, J., concurring in the judgment).

115. Paul B. Stephan III, The First Amendment and Content Discrimination, 68 Va. L. Rev. 203, 238 n.155 (1982) (providing "nonexhaustive list of the factors one might consider in weighing a particular speaker's claim to speak in a particular place").

assesses whether the restriction is narrowly tailored to serve an important government interest, and whether the restriction leaves open ample alternative channels for communication of the message.[116] These tests balance the harm from the speech against the harm to the speech, including whether there are other alternatives for voicing a particular message.

In the context of speech by government employees, the Court has held that the First Amendment protects government employees acting in their capacity as citizens speaking on matters of public concern, and courts uphold "only those speech restrictions that are necessary for their employers to operate efficiently and effectively."[117] When considering speech by students in public schools, the Supreme Court protects high-value speech and gives schools leeway to regulate speech it considers harmful. For example, the Court has held that schools may not restrict speech simply because it is "offensive," because that might pertain to political or religious speech, but can punish students for speech that is "reasonably viewed as promoting illegal drug use."[118] The Supreme Court's approach to restrictions that are not content-based or do not target the general public is even more explicit in weighing the burdens of the speech to the government or to the citizenry against the burdens to the speech.

4. *The Means of Regulating Speech.* — The Court's evaluation of the way in which the government regulates speech, including the severity of the penalty and the form of restriction, often reflects free speech consequentialism, even if the Court does not explicitly or consistently marshal consequentialist tests.

Penalty sensitivity, or attention to the severity of the penalty in assessing whether a speech restriction is constitutional, often happens without recognition in the First Amendment context.[119] If a law's greater coercive power, for example criminal penalties instead of a civil fine (or larger versus smaller fines),[120] translates to less leeway to restrict speech, this would reflect free speech consequentialism in the methods of regulating speech. In that case, the greater the coercive effect of the law, the greater the burden on speech, the more justification the regulation

116. See McCullen v. Coakley, 134 S. Ct. 2518, 2529 (2014) (describing "somewhat wider leeway" given to speech regulations unrelated to content).

117. Garcetti v. Ceballos, 547 U.S. 410, 419 (2006).

118. Morse v. Frederick, 551 U.S. 393, 409 (2007).

119. See Michael Coenen, Of Speech and Sanctions: Toward a Penalty-Sensitive Approach to the First Amendment, 112 Colum. L. Rev. 991, 994 (2012) ("In particular, the *severity* of the penalty imposed—though of central importance to the speaker who bears it—does not normally affect the merits of his free speech claim.").

120. See Wendy Gerwick Couture, White Collar Crime's Gray Area: The Anomaly of Criminalizing Conduct Not Civilly Actionable, 72 Alb. L. Rev. 1, 42–43 (2009) ("This relationship between civil and criminal liability is supported by the general rationale that criminal sanctions are more severe than civil liability.").

requires. Courts explicitly maintain that it is generally "immaterial for First Amendment purposes whether speech is suppressed under the criminal law or by 'penalties' imposed by tort law."[121] The Supreme Court has noted that, like the application of criminal law, tort law's penalties constitute state action and are therefore equally subject to First Amendment restrictions.[122] Yet, there are notable examples of penalty-sensitivity in free speech jurisprudence. For example, the Supreme Court recognizes that "a law imposing criminal penalties on protected speech is a stark example of speech suppression."[123] The doctrine of defamation, with its tailoring of penalties depending on whether speech reflects a matter of public concern, and whether the plaintiff can prove actual malice,[124] is also an example of penalty sensitivity. The Court is also less wary of vague or overbroad laws that restrict speech in the civil context than in the criminal context.[125] And Congress has more leeway in conditioning funding on content-based determinations than in regulating the content of speech via criminal or tort law,[126] although this discrepancy stems in part from the government's role in creating the speech and allocating its own funds.[127]

Further, the timing of the restraint matters, perhaps for free speech consequentialist reasons. Prior restraints on speech, such as preliminary injunctions, are afforded much less deference than ex post penalties, and come to the Court "with a heavy presumption against . . . constitutional validity."[128] This doctrinal choice may embody free speech consequentialism in its overriding concern for the burdens prior restraints

---

121. Yakubowicz v. Paramount Pictures Corp., 536 N.E.2d 1067, 1071 (Mass. 1989).

122. See N.Y. Times Co. v. Sullivan, 376 U.S. 254, 265 (1964) ("[T]he Alabama courts have applied a state rule of law which petitioners claim to impose invalid restrictions on their constitutional freedoms of speech and press. . . . The test is not the form in which state power has been applied but . . . whether such power has in fact been exercised.").

123. Ashcroft v. Free Speech Coal., 535 U.S. 234, 242–44 (2002) (striking down ban on criminal statute that criminalized "virtual child pornography" not involving actual minors).

124. See Sacks, supra note 15, at 1109–11 (discussing availability of different penalties, from presumed damages to actual damages to punitive damages, in defamation lawsuits depending on whether plaintiff is public or private figure and whether speech is matter of public concern).

125. See Coenen, supra note 119, at 1010–11 (noting Court's "'greater tolerance of enactments with civil rather than criminal penalties'" (quoting Vill. of Hoffman Estates v. Flipside, Hoffman Estates, Inc., 455 U.S. 489, 499 (1982))).

126. See Nat'l Endowment for the Arts v. Finley, 524 U.S. 569, 587–88 (1998) ("Finally, although the First Amendment certainly has application in the subsidy context, we note that the Government may allocate competitive funding according to criteria that would be impermissible were direct regulation of speech or a criminal penalty at stake.").

127. See id. at 587–89 ("There is a basic difference between direct state interference with a protected activity and state encouragement of an alternative activity consonant with legislative policy." (citing Maher v. Roe, 432 U. S. 464, 475 (1977)).

128. Org. for a Better Austin v. Keefe, 402 U.S. 415, 419 (1971) (internal quotation marks omitted) (quoting Carroll v. President of Princess Anne, 393 U.S. 175, 181 (1968); Bantam Books, Inc. v. Sullivan, 372 U.S. 58, 70 (1963)).

place on speech, although this choice may be based on the history of the First Amendment and its response to prior restraints on the press in England.[129]

In sum, in the categorization of different types of speech, the scrutiny that attaches to different categories, and the methods of regulating speech, free speech consequentialism pervades much of First Amendment doctrine. The next section shows that, despite its problematic nature, free speech consequentialism is both unavoidable and, in some form, normatively desirable.

C. *The Futility of Alternative Approaches*

In resisting calls for free speech consequentialism, some courts and scholars have touted alternative ways of determining which speech must be protected and which speech can be regulated. Ultimately, however, these approaches all lead back to free speech consequentialism. Without any harms balancing at some stage in the inquiry, First Amendment doctrine would be absolutist in ways that yield unpalatable results. That is, in addition to being ubiquitous, some version of free speech consequentialism has thus far proven both unavoidable and normatively necessary in protecting speech values.

In recent years, the anticonsequentialist account that has gained the most momentum is the view that government motive is a primary factor in determining the constitutionality of a challenged regulation.[130] This purposivist view, advanced by then-Professor, now-Justice Elena Kagan, claims that First Amendment doctrine was established to discover impermissible governmental motive.[131] Purposivism explains why strict scrutiny usually attaches to content-based regulations and deferential review attaches to generally applicable laws that impact speech, because greater justifications are needed the more likely it appears that the government is suppressing speech based on its content or viewpoint.[132] Further, purposivism

---

129. See Times Film Corp. v. City of Chicago, 365 U.S. 43, 63–64 (1961) ("'[T]he history . . . indicates that a major purpose of the First Amendment guaranty of a free press was to prevent prior restraints upon publication, although . . . the liberty of the press is not limited to that protection.'" (quoting Joseph Burstyn, Inc. v. Wilson, 343 U.S. 495, 503 (1952))).

130. See Elena Kagan, Private Speech, Public Purpose: The Role of Governmental Motive in First Amendment Doctrine, 63 U. Chi. L. Rev. 413, 422 (1996) (noting impermissible governmental motive would be, for example, "purely censorial—a simple desire to blot out ideas of which the government or a majority of its citizens disapproved").

131. See id. at 413–14 ("First Amendment law, as developed by the Supreme Court over the past several decades, has as its primary, though unstated, object the discovery of improper governmental motives."); see also Jed Rubenfeld, The First Amendment's Purpose, 53 Stan. L. Rev. 767, 794 (2001) (explaining narrow-tailoring test in strict scrutiny is used to "smok[e]-out" unconstitutional purpose).

132. See Rosenthal, supra note 1, at 2 ("[T]he Court's use of strict scrutiny for laws directed at the content of speech is . . . based on the risk that content regulation reflects

explains why tests have developed to limit unfettered discretion of officials when granting permits or licenses.[133] For purposivists, "[a] jurisprudence based on . . . the likelihood of an illicit governmental motive is thought to be preferable to one that endeavors to balance the costs and benefits of a challenged regulation of speech because the latter inquiry . . . amount[s] to an invitation for unbridled judicial activism."[134]

Free speech purposivists offer an important descriptive account of free speech jurisprudence, and a significant number of scholars believe it should be adopted normatively.[135] However, as a descriptive matter, the Court's adoption of purposivism has been inconsistent. Further, although the principled nature and simple coherence of purposivism has great normative appeal, pure purposivism without any consequentialism requires an absolutist approach to free speech that presents many problems.

Descriptively, the Court itself has noted that "our cases have consistently held that '[i]llicit legislative intent is not the *sine qua non* of a violation of the First Amendment.' . . . 'We have long recognized that even regulations aimed at proper governmental concerns can restrict unduly the exercise of rights protected by the First Amendment.'"[136] Perhaps, then, illicit governmental intent is *sufficient*, even if not necessary, to invalidate a law. However, even this notion is belied by the Court's fairly expansive view of content neutrality, even in the face of governmental targeting of particular speech. In *Hill v. Colorado*, for example, the Court upheld a law, inspired by the activities of abortion protesters, limiting protestors' access to individuals entering health facilities.[137] The Court noted that statutes may be categorized as content-neutral and subject to deferential scrutiny even if they are motivated by the activities of a particularly unpopular group.[138] The Court elaborated that a statute prohibiting solicitation in airports, despite being motivated by activities of Hare Krishnas, is not content-based, nor is a statute banning sitting at a lunch counter without ordering food, even if "enacted by a racist legislature that hated civil rights protesters (although it might raise separate questions about the State's legitimate interest at issue)."[139]

---

official hostility to disfavored content, coupled with the difficulties that inhere in requiring those seeking to vindicate First Amendment rights to prove illicit motivation.").

133. See id. at 2–3 ("[T]he insistence of First Amendment jurisprudence on adequate standards to cabin the discretion of officials who regulate speech is said to be based on a concern that absent such standards, regulation will be infected by difficult-to-detect illicit motives.").

134. Id. at 3.

135. See Rubenfeld, supra note 131, at 787–93 (providing account of underlying rationales of purposivism).

136. Simon & Schuster, Inc. v. Members of the N.Y. State Crime Victims Bd., 502 U.S. 105, 117 (1991) (quoting Minneapolis Star & Tribune Co. v. Minnesota Comm'r of Revenue, 460 U.S. 575, 592 (1983)).

137. 530 U.S. 703, 707–08, 734–35 (2000).

138. Id. at 723–24.

139. Id. at 724.

This half-hearted application of purposivism continued in *McCullen v. Coakley*, another case involving the "buffer zone" required between protestors (or anyone standing on a public sidewalk near a facility that performs abortions) and individuals entering clinics that perform abortions.[140] Although the Court invalidated the thirty-five foot buffer zone in *Coakley* as too large, and thus too speech-restrictive, the majority held that the regulation was content-neutral because the government's goal was to prevent patients from being intimidated.[141] As Justice Scalia's concurrence noted, "It blinks reality to say . . . that a blanket prohibition on the use of streets and sidewalks where speech on only one politically controversial topic is likely to occur—and where that speech can most effectively be communicated—is not content based."[142]

Although the Court may not fully embrace purposivism, the free speech purposivists have articulated a fairly compelling descriptive account of some of First Amendment jurisprudence, including the different types of scrutiny corresponding to the likelihood that the regulation at issue reflects impermissible motives.[143] However, at some point, the Court inevitably considers both the value of speech and the harm it causes, not just the government's impermissible motive.

Scholars who have proposed accounts that entail some form of purposivism still must fall back on some degree of free speech consequentialism.[144] In a recent article, Genevieve Lakier advocates for a different form of purposivism that entails a strong form of free speech consequentialism.[145] Lakier denounces the current Court's use of history in categorizing speech and maintains that the value of speech should be determined in light of the purposes of the First Amendment, for example whether the speech relates to "matters of public concern."[146] However, this type of purposivism closely resembles explicit consequentialist views about the First Amendment, as Lakier acknowledges that her view "vest[s] judges with considerable discretion to determine when a particular category of speech does or does not advance the First Amendment's purposes."[147]

Lakier argues that courts should make substantive decisions about which speech it wants to protect, regardless of history, by challenging the view that our currently unprotected categories such as obscenity and libel

---

140. 134 S. Ct. 2518, 2525 (2014).

141. Id. at 2531.

142. Id. at 2543 (Scalia, J., concurring in the judgment).

143. See supra notes 131–132 and accompanying text (discussing justifications for strict scrutiny).

144. See supra notes 98–102 and accompanying text (discussing how doctrinal structure of many First Amendment tests also reflects free speech consequentialism).

145. Genevieve Lakier, The Invention of Low-Value Speech, 128 Harv. L. Rev. 2166 (2015).

146. Id. at 2225–26, 2228–29.

147. Id. at 2227.

were historically unprotected.[148] Lakier's account often disproves itself, however, as many of the categories she describes received almost the same protections and definitions as exist today. For example, fighting words could be regulated only when they were likely to incite a particular person to violence,[149] just as today. Further, Lakier's explanations of how libel and obscenity, currently unprotected speech, did receive some historical protection actually show that, in the eighteenth and nineteenth centuries, as it is now, truthful speech was protected as nonlibelous and obscenity was prosecuted at common law at the time of the founding.[150] Lakier's purposivist view of the First Amendment would eliminate history as a nonconsequentialist way of crafting First Amendment doctrine, which paves the way for more instrumental thinking about when speech is and is not valuable, and would give the government more leeway in regulating speech-based harm.[151]

A purely purposivist account, by contrast, would be unpalatable even to the most anticonsequentialist Justices. As Jed Rubenfeld argues, "A purposivist view of the First Amendment does not involve *balancing*. It is absolute. It does not purport to determine if the constitutional 'costs' in a given case are 'outweighed' or 'justified' by the governmental 'benefits.'"[152] Descriptively, this is a huge departure from current First Amendment doctrine,[153] and normatively, although simpler, pure purposivism will both overprotect and underprotect speech.

Pure purposivism would invalidate all regulations that manifest, or pose a risk of, impermissible governmental censorship, regardless of the magnitude of the governmental interest or the narrowness of the statute. Because it is often difficult to discern government motive, the tests applied to serve as proxies for impermissible government motive would sometimes invalidate regulations not created with a desire to censor particular viewpoints. Employing some version of pure purposivism, such as a categorical approach designed to track impermissible content-based motives, would require courts to simply ignore, or severely minimize, the harms that flow from speech. (This is not true of our current jurisprudence, and most purposivist scholars do not advocate for pure purposivism without any of what this Article deems free speech conse-

148. See id. at 2168 (arguing distinction between high- and low-value speech actually arose during New Deal era).

149. Id. at 2190–91 (explaining that, even in 1800s, "courts made clear that there were limits on the government's ability to criminally punish speech merely because of its offensive or insulting content").

150. Id. at 2184–86, 2206 (demonstrating truth was historically defense to libel and that obscenity could be prosecuted).

151. Id. at 2223–26 (arguing "matter of public concern [test] would also provide courts the flexibility to recognize novel categories of low-value speech, even when these kinds of speech either did not exist in the eighteenth or nineteenth century").

152. Rubenfeld, supra note 131, at 779.

153. See id. at 778 (acknowledging proposed approach "calls for a reconceptualization of the basic structure of free speech law").

quentialism.) To avoid any incorporation of free speech consequentialism, courts could simply define away harms that flow from speech as not actual harms, or harms society must bear.[154] Courts could then focus solely on whether speech fits into an unprotected category, and protect all other speech absolutely. Or, courts could use some other metric to inquire directly about governmental motivation, and invalidate absolutely any statutes that failed their test. Pure purposivism would be a radical approach, because the government's interest in regulating speech could not be considered at all.

More alarmingly, pure purposivism would underprotect speech. Absent risk of improper governmental motive, regulations that are extremely burdensome to speech should still be upheld. Thus, if a regulation targeted only the nonspeech elements of a particular activity but actually infringed upon a great deal of speech, the regulation would be perfectly valid. A law that prohibited walking near health facilities during daylight hours unless seeking medical care, to use the abortion buffer zone example mentioned above,[155] would be perfectly legitimate even if the buffer zone were very large, and even if important speech activities occurred on those public walkways, so long as the regulation did not evidence an impermissible government motive. And, because it is often difficult to discern government motive, the tests applied to ferret out government motive would likely miss impermissible motive in many cases. Every content-neutral regulation or law targeting expressive conduct, like nudity or flag burning, would likely be upheld, unless impermissible motive were obvious (or unless courts had some way of divining impermissible motive) because the laws on their face do not target expression or reflect an impermissible governmental motive.[156] Currently, even content-neutral laws, or laws aimed at conduct that negatively impacts speech, still receive some form of First Amendment scrutiny.[157]

Further, under a strong version of purposivism that inquires only into government motive, laws that target the secondary effects of speech, like violence or diminishing property values,[158] could be upheld with no further balancing or examination of the impact on speech, even if these

---

154. Although the punishment of murder is recognized as a harm, generally those deterred from committing murder are not considered to be harmed, even though their preference for killing has been subverted.

155. See supra notes 139–140 and accompanying text (noting Court's upholding of regulation to turn on content-neutrality even where buffer zone was found to be too large).

156. See Rubenfeld, supra note 131, at 780 (outlining Posner's bullfighting example, noting torturing of bulls can only be banned without reference to what it communicates, despite expressive components).

157. See United States v. O'Brien, 391 U.S. 367, 366–77 (1968) (laying out requirements for "sufficiently justified" government regulation involving "speech" and "nonspeech" elements).

158. See Stone, Content-Neutral Restrictions, supra note 18, at 115 (discussing case where court characterized content-based ordinance as content-neutral because of "secondary effects").

laws directly target speech based on its content. For example, certain song lyrics could be censored based on their propensity to lead to violence. With no impermissible censorial governmental motive, simply a legitimate government desire to stop violence, this type of regulation could be upheld in a purposivist universe but likely would be invalidated when subject to the current tests ensuring that even content-neutral regulations do not unduly suppress speech.

Employing some form of free speech consequentialism is thus a necessary evil. Descriptively, free speech consequentialism pervades the doctrine and is employed by even those who eschew consequentialism. In addition, some amount of free speech consequentialism is normatively beneficial because any alternative to balancing would involve absolutes that would both overprotect and underprotect speech.

D. *Weighing the Different Types of Free Speech Consequentialism*

The conclusion that some degree of free speech consequentialism is irresistible does not mean that all free speech consequentialism is equal. Put another way, to say that everyone is a free speech consequentialist is not to say that all free speech jurisprudence is equally arbitrary, unprincipled, or political. This section explores the virtues and vices of employing free speech consequentialism at various stages, or levels of abstraction, within First Amendment doctrine.

As John Hart Ely notes, an "absolutist" categorization approach and a balancing approach to the First Amendment "are more helpfully employed in tandem, each with its own legitimate and indispensable role in protecting expression."[159] Ely posits that a categorical approach is proper when laws target the message expressed, and a balancing approach is appropriate "where the evil the state would avert does not grow out of the message being communicated."[160] However, as decisions like *Alvarez* indicate, even cases involving self-professed categorizers involve some, if weakened, form of free speech consequentialism. The important choice, therefore, is between different types of free speech consequentialism.[161]

It may be true that "the decision to draw a line between protected and regulated speech will always be a decision to advance some interests and discourage others, will always, that is, be a political decision."[162] Although First Amendment doctrine may be politics all the way down,

---

159. John Hart Ely, Flag Desecration: A Case Study in the Roles of Categorization and Balancing in First Amendment Analysis, 88 Harv. L. Rev. 1482, 1501 (1975); see also Tushnet, Anti-Formalism, supra note 1, at 1531 (noting many agree "law of the first amendment ought to consist of a mix of relatively clear rules, used in some areas, and balancing, used in others").

160. Ely, supra note 159, at 1501.

161. See supra notes 22–27 and accompanying text (discussing *Alvarez* and concluding Court adopted some form of free speech consequentialism).

162. Fish, supra note 4, at 1062.

different approaches to free speech consequentialism are differently political. Free speech doctrine attempts, although does not always succeed, to honor "the basic neutrality of First Amendment law" by "offer[ing] judges little opportunity to read their own preferences and ideological commitments into the Constitution."[163]

The more abstract the inquiry and the weaker the incorporation of free speech consequentialism, the more likely First Amendment jurisprudence can be based on general principles that ultimately may lead jurists to adopt results they do not personally favor in any given case. However, adopting free speech consequentialism at this wholesale level also has its pitfalls, and different approaches to free speech consequentialism have different virtues and vices. Free speech consequentialism at the retail level, where courts first determine the value of the speech on a sliding scale and then subject it to various degrees of constitutional scrutiny depending on that valuation, involves woefully ad hoc determinations. But free speech consequentialism performed at the wholesale level, to determine which categories of speech receive protection and what goes into those categories, risks suppressing entire categories of speech based on subjective judgments.[164]

The most robust incorporation of free speech consequentialism was the retail approach taken by Justice Breyer's concurrence in *United States v. Alvarez*,[165] which shares substantial similarities with the concurrence in *Brown v. Entertainment Merchants Association*,[166] even though different Justices were involved. In *Alvarez*, the plurality applied strict scrutiny to invalidate a federal statute banning lies about receiving a military honor; this rigorous constitutional scrutiny was based on the idea that lies are not historically unprotected and are therefore high-value speech.[167] In contrast, Justice Breyer's concurrence reflects the consequentialist determination that false speech has lesser value than truthful speech but still

---

163. Lakier, supra note 145, at 2175. Lakier argues that courts' neutrality is just a veneer, and that courts more transparently import substantive value judgments into the First Amendment. Id. at 2225–26.

164. Because balancing interests likely requires judges to "adopt some form of utilitarianism," assessing interests at various levels of abstraction "has obvious affinities to the philosophers' distinction between act utilitarianism and rule utilitarianism, and raises problems of a sort that philosophers have discussed in that context." Tushnet, Anti-Formalism, supra note 1, at 1511–12.

165. 132 S. Ct. 1537, 2551–56 (2012) (Breyer, J., concurring in the judgment). In *Alvarez*, the defendant had lied about the receipt of the Congressional Medal of Honor. See supra note 104 and accompanying text (describing factual background of *Alvarez*).

166. 131 S. Ct. 2729 (2011). In *Brown*, Justice Alito and Chief Justice Roberts concurred to say that they would have applied intermediate scrutiny to California law restricting the sale of violent video games to minors. Id. at 2747–49 (Alito, J., concurring in the judgment).

167. See *Alvarez*, 132 S. Ct. at 2544 (plurality opinion) (Kennedy, J.) (noting "common understanding that some false statements are inevitable if there is to be an open and vigorous expression of views in public and private conversation, expression the First Amendment seeks to guarantee").

produces important benefits.[168] This led Justice Breyer to apply an intermediate level of constitutional scrutiny, which in turn, in a second, and different, version of consequentialism, assessed the means and ends of the statute to ensure proportionality to the interest served.[169] Ultimately, Justice Breyer concluded that the government's goal of preventing deceptions involving military honors could have been achieved in a manner that was substantially less burdensome to speech.[170]

The problems with this approach are myriad. First, consideration of the benefits of speech under this robust form of free speech consequentialism requires deciding whether the benefits of speech are (1) only those that advance the perceived instrumental goals of the First Amendment, or (2) any consequences of the speech that increase happiness or another virtue that is broader than the perceived aim of the First Amendment. Justice Breyer articulated the benefits of deception as lubricating social interactions, protecting privacy, avoiding embarrassment, and promoting thought that ultimately leads to truth.[171] Some scholars, however, consider only the instrumental benefits of speech that fulfill their perceived vision of the goals of the First Amendment[172]—this is both a narrower and more explicit (and perhaps more dangerous and easily manipulated) form of free speech consequentialism. The plurality's approach avoids this choice in how to weigh speech's benefits by assuming that almost all protected speech is high value and then examining the harms.

Further, Justice Breyer selected the level of generality with which he would consider the speech as false statements generally, but could just as easily have selected lies about awards or honors received more specifically. This approach could have led to either applying more deferential constitutional scrutiny or a different result under intermediate review. Moreover, although Justice Breyer ultimately concurred in the judgment, an inquiry into whether "it is possible substantially to achieve the Government's objective in less burdensome ways"[173] is indeterminate. Justice Breyer suggested that the government could require knowledge of falsity, but claimed that even a mens rea requirement would not diminish "other

168. See id. at 2553 (Breyer, J., concurring in the judgment) ("I must concede, as the Government points out, that this Court has frequently said or implied that false factual statements enjoy little First Amendment protection.").

169. See id. at 2551 (Breyer, J., concurring in the judgment) ("In determining whether a statute violates the First Amendment, this Court has often found it appropriate to examine the fit between statutory ends and means.").

170. See id. at 2556 (Breyer, J., concurring in the judgment) (noting how government could have required proof of specific or material harm as example).

171. See id. at 2553 (Breyer, J., concurring in the judgment) (listing benefits of deception in society).

172. See, e.g., Chen & Marceau, supra note 7, at 1437 ("[S]ome lies—what we call *high value lies*—have instrumental value that advances the goals underlying freedom of speech.").

173. *Alvarez*, 132 S. Ct. at 2555 (Breyer, J., concurring in the judgment).

First Amendment risks, primarily risks flowing from breadth of coverage."[174] According to Justice Breyer, the Stolen Valor Act, which he determined to have a substantial justification, perhaps would have been saved if it required a showing of specific harm, or material harm, or focused on more harmful deception involving particularly exalted military honors.[175]

These decisions regarding the application of intermediate scrutiny appear rather arbitrary. Fashioning a more "finely tailored"[176] statute, following Justice Breyer's guidance, may even still end up chilling a large amount of speech that Justice Breyer acknowledges is beneficial.[177] Ultimately, Justice Breyer's determination of when speech can be regulated depends on his idiosyncratic calculation of when the benefits of this speech are outweighed by their harms. Justices with a more protective attitude toward speech will make this calculation differently, and Justices who believe speech is less valuable than currently perceived will be more amenable to being convinced that lesser harms render speech punishable.[178]

Perhaps incorporating free speech consequentialism at a higher level of abstraction would remove some of the ad hoc nature of the inquiry. At the highest levels of abstraction, free speech consequentialism may not greatly undermine our aspiration that free speech jurisprudence be politically neutral. For example, imagine a statute that outlaws criticizing a member of Congress for having an extramarital affair. Perhaps legislators believe that this type of accusation is too pervasive in political discourse and is not a legitimate concern of the electorate. Adopting, as a starting point, a categorical or purposivist approach would immediately assess this law as a content-based regulation of protected speech, and likely impermissibly motivated. No matter how much members of the Court might like members of Congress, and maybe even like one member in particular who has come under a particularly obnoxious and virulent

---

174. Id. (Breyer, J., concurring in the judgment).

175. See id. at 2555–56 (Breyer, J., concurring in the judgment) (stating First Amendment risks could be diminished if statute "insist[ed] upon a showing that the false statement caused specific harm or at least was material, or focus[ed] its coverage on lies most likely to be harmful").

176. Id. (Breyer, J., concurring in the judgment).

177. See id. at 2556 (Breyer, J., concurring in the judgment) ("In the political arena a false statement is more likely to make a behavioral difference . . . but at the same time criminal prosecution is particularly dangerous . . . and consequently can more easily result in censorship of speakers and their ideas.").

178. Balancing in constitutional law generally is susceptible to these same criticisms, but courts should be extra cautious about creating indeterminacy in free speech jurisprudence because of the chilling effects. See Melville B. Nimmer, Does Copyright Abridge the First Amendment Guarantees of Free Speech and Press?, 17 UCLA L. Rev. 1180, 1183–84 (1970) (discussing dangers of ad hoc balancing in creating chilling effect on free speech).

attack, the Court will have a tough time justifying this regulation under strict scrutiny.

Of course, the Court would also have a tough time justifying this regulation under Justice Breyer's approach, balancing the costs and benefits on a more case-by-case basis, but there is more wiggle room for subjectivity. The lower the level of generality at which the speech is considered, and the later the stage of jurisprudence at which balancing is incorporated, the easier it is for a jurist's political views and biases to enter the equation. At the extremes, if "political speech" or "speech critical of Congress" is considered, a court must examine these categories abstractly. Conversely, if speech critical of a particular Senator that is largely unsubstantiated but not demonstrably false is considered, courts can begin to allow their views about adultery, Congress, and even that Senator to enter the equation. It may even be impossible to avoid having views on these subjects enter the equation, if a court is considering the costs and the benefits of that particular speech at a retail level.

It is true that categorically deciding that, for example, obscenity is unprotected, favors the interests of more prudish individuals over the interests of those who enjoy prurient speech. But there is less room at this high level of abstraction for discrimination of particular viewpoints in the categorical and purposivist approaches. In *United States v. Stevens*, for example, where the Court invalidated a law prohibiting the depiction of animal cruelty, likely all of the Justices objected to the speech—pornography made by torturing animals—but Chief Justice Roberts's categorical approach chastened their ability to incorporate those biases.[179] If speech depicting violent torture of animals could be outlawed, the Justices reasoned in oral argument, what about videos of human sacrifice, or activities legal in other countries but illegal in the United States?[180] Thinking of the speech at issue in *Stevens* as "speech" at the highest levels of abstraction and generality permits the least subjectivity. More and more subjectivity enters the equation if we consider speech depicting illegal activity, speech depicting torture, speech depicting torture of animals, or speech depicting torture of animals for prurient reasons.[181] However, even at the highest levels of abstraction or generality, balancing is always

179. See 559 U.S. 460, 481–82 (2010) (holding statute overly broad without deciding its constitutionality as applied to extreme animal cruelty cases).

180. See Transcript of Oral Argument at 30, 45–46, *Stevens*, 559 U.S. 460 (No. 08-769) ("[T]hen what about people who—who like to see human sacrifices? Suppose that is legally taking place someplace in the world. I mean, people here would probably love to see it. Live, pay per view, you know, on the human sacrifice channel.").

181. Of course, this type of speech can be regulated if restricted to obscene speech of this type. See United States v. Richards, 755 F.3d 269, 275–76 (5th Cir. 2014) (upholding federal statute crafted in response to *Stevens* because it incorporates *Miller* obscenity test and proscribes only unprotected speech).

unavoidably subjective, and even political, in that it prioritizes certain values, ideals, and experiences over others.[182]

All normative judgments, including ones like "murder is wrong," are in a sense subjective, and judges must make normative judgments. But, in the First Amendment context, allowing judges to determine which speech is protected for ideological reasons or which speech harms bother them personally would give courts the power to do something legislators cannot. Courts would then become the censors instead of the government. As a threshold matter, then, when faced with a content-based restriction on speech, courts should use a categorical approach, with a strong default against creating new categories of unprotected speech,[183] and also craft jurisprudence that is sensitive to purposivist concerns.[184] Although it is possible to weigh the costs and benefits of speech without letting personal biases enter the equation, human beings are subject to cognitive biases that make this very difficult.[185] A jurisprudence that, on a case-by-case basis, weighs the costs and benefits of speech is, as Chief Justice Roberts notes, "startling and dangerous."[186] Even creating new categories of unprotected speech based on this weighing at a higher level of abstraction would, as a general matter, allow wide swaths of speech to be suppressed based on the subjective weighing of nine Justices regarding which types of speech are valuable.

To minimize the dangers of free speech consequentialism, courts could begin with the premise that unprotected categories of speech are basically set, or derived from a historical account of the First Amendment. Courts could also begin with the premise that strict scrutiny, applied to content-based restrictions for protected categories of speech, is quite strict, and usually "fatal in fact."[187] However, new technologies will create new categories of speech and new types of harms, some of which courts may wish to regulate, even at the wholesale level. And balancing at the

---

182. See supra note 162 and accompanying text (noting political aspect of balancing).

183. See infra notes 275–282 and accompanying text (explaining why creating new categories of unprotected speech should be disfavored).

184. See supra notes 130–135 and accompanying text (describing merits of purposivist approach).

185. See Dan M. Kahan et al., "They Saw a Protest": Cognitive Illiberalism and the Speech–Conduct Distinction, 64 Stan. L. Rev. 851, 883–84 (2012) (showing study subjects find speech they disagree with more threatening and are more likely to view such demonstrators as obstructing pedestrian traffic).

186. United States v. Stevens, 559 U.S. 460, 470 (2010).

187. See Gerald Gunther, The Supreme Court, 1971 Term—Foreword: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection, 86 Harv. L. Rev. 1, 8 (1972) (describing strict scrutiny inquiry applying to violations of fundamental rights, including free speech liberties, as "'strict' in theory and fatal in fact"). Recently, however, scholars have begun to challenge the idea "that strict scrutiny is an outcome determinative, always (or nearly always) fatal test." Adam Winkler, Fatal in Theory and Strict in Fact: An Empirical Analysis of Strict Scrutiny in the Federal Courts, 59 Vand. L. Rev. 793, 808 (2006).

retail level needs to remain an option as well, or else courts may decline to extend First Amendment protections into new contexts if they fear the jurisprudence will be too absolute. Given that free speech consequentialism is unavoidable, what should courts do to preserve a principled, speech-protective First Amendment regime? There must be a better way to marshal flexibility but constrain the negative effects of balancing.

II. MODIFIED FREE SPEECH CONSEQUENTIALISM

Free speech consequentialism is a necessary and accepted part of First Amendment jurisprudence. Yet existing proposals to discipline and incorporate free speech consequentialism are wanting. A constraint on free speech consequentialism would avoid many of the pitfalls described in the previous section: Courts should engage in balancing speech harms against the benefits of speech only where those harms are analogous to harms that flow from conduct, based on the properties that distinguish classic instances of speech (such as political satire or protest) from paradigmatic cases of conduct (such as battery or theft). This limiting principle focuses our energies on the right questions. There are strong reasons why harms that flow uniquely from speech should not be balanced against the benefits of speech. This Part outlines the proposal, provides the justifications, and confronts various objections.[188]

A. *Limiting the Types of Harms Considered*

In cases where it is unavoidable, courts should incorporate free speech consequentialism in a way that avoids undermining free speech principles or enabling viewpoint discrimination on the part of legislators, executives, or judges. To do so, courts should distinguish between the types of harms specific to speech and the types of harms generally caused by conduct, which are typically harms that directly impact material, concrete interests in the physical world. As an example, anti-spam legislation can be likened metaphorically to prohibitions on trespass to chattels.[189] Threatening to inflict bodily harm, even without doing so, is also

188. This Article does not devote as much time to the proper consideration of speech's benefits—i.e., whether courts should consider benefits related only to the purposes of the First Amendment or all benefits to individuals or society emanating from free expression of the speech—when resorting to free speech consequentialism. See supra section I.D (giving brief overview of different benefits considerations). However, as section I.D suggests, generally using a categorical approach where most speech is considered high value requires less emphasis on the benefits of speech. Perhaps courts should consider all of the benefits of speech to be less narrowly consequentialist.

189. Some scholars argue that spam does not cause an actual trespass to chattels, because the electrons from spam that "invade" a recipient's server or inbox do no damage to the physical hardware. See Michael J. Madison, Rights of Access and the Shape of the Internet, 44 B.C. L. Rev. 433, 467–68 (2003). However, because spam does occupy bandwidth, courts have upheld anti-spam lawsuits on the theory that spam "intermeddles" with computers or servers such that they are impaired or harmed, or the owner is deprived of

unprotected because the speech actually represents the direct specter of imminent unlawful conduct.[190] Fraud is tied to the harm of undermining people's choice relating to their property or pecuniary interests.[191] In conducting its consequentialist calculus, when defining new unprotected categories of speech, determining whether speech fits into a particular category, or applying the constitutional scrutiny that corresponds to particular categories, a court should give weight to the harms caused by speech only when these harms can be analogized to conduct harms.

This proposal depends on the difference between harms typically caused by speech and harms typically caused by conduct. Setting aside the values and benefits robust free speech protections serve, classic "speech harms," or harms arising from speech that is generally considered protected, are less amenable to regulation. Speech harms are generally context dependent or bound up in our values, perceptions, and identity.[192] Speech harms are either largely emotional in nature, and thus subjectively felt,[193] or caused by diffuse parties or intermediaries, not primarily the speaker herself.[194] Speech harms involve both what we, as

---

their use for a substantial period of time. See, e.g., CompuServe Inc. v. Cyber Promotions, Inc., 962 F. Supp. 1015, 1027–28 (S.D. Ohio 1997) (granting anti-spam preliminary injunction based on theory of trespass to chattels).

190. See Edward J. Eberle, Cross Burning, Hate Speech, and Free Speech in America, 36 Ariz. St. L.J. 953, 993 (2004) ("The close association of threats and intimidation with action is what makes them proscribable."); see also Virginia v. Black, 538 U.S. 343, 359 (2003) ("'True threats' encompass those statements where the speaker means to communicate a serious expression of an intent to commit an act of unlawful violence to a particular individual or group of individuals.").

191. See, e.g., 18 U.S.C. § 1341 (2012) (defining mail fraud as involving "any scheme or artifice to defraud, or for obtaining money or property by means of false or fraudulent pretenses, representations, or promises"); John C.P. Goldberg, Anthony J. Sebok, & Benjamin C. Zipursky, The Place of Reliance in Fraud, 48 Ariz. L. Rev. 1001, 1011 (2006) (arguing core of legal wrong of fraud is interfering with individual's ability to make "certain kinds of choices in certain settings free from certain forms of misinformation").

192. See Schauer, Harm(s) and the First Amendment, supra note 7, at 100–01 (explaining how some harmful words cause injury because of "effect that the words are expected to have on the mental state of those who hear (or see) them under certain circumstances").

193. This would include Schauer's harms of verbal assault. See id. (describing category of "harms of verbal assault").

194. These types of harms would fall within Schauer's harms of advocacy, or Han's dissemination-based harm. See supra notes 37–45 and accompanying text (describing Schauer's and Han's framework). Participant harms, like animal crush videos or child pornography, can also be placed in this category. The speaker's conduct in producing the video causes the primary harm, but the harm from the speech is then caused when others watch the video or are compelled themselves to produce more of this type of speech. See Han, supra note 7, at 1661 (categorizing harm to animals in "crush videos" and harm to children in child pornography as harms formed by both creation and dissemination of speech); Schauer, Harm(s) and the First Amendment, supra note 7, at 103 (categorizing animal crush videos and child pornography videos as dangerous due to "participant harms"). Or, if the harm comes from knowing the speech exists, the speech would be

listeners, value and how others in society react. Finally, the harm caused by speech may in fact be a benefit, as speech can serve as an agent of change that actually alters what we perceive to be harmful and beneficial without inflicting physical damage.[195]

These characteristics of speech explain why jokes that were once considered funny may come to be considered offensive, or vice versa. As another example, pornography, which is protected speech unless it rises to the level of obscenity, is said to cause harm by transmitting an idea about the subjugation of women,[196] and this idea then is spread by diffuse parties. As Schauer notes, "words cause an injury not because of something they necessarily do in all contexts, and not because of some intrinsic property of words as words, but because of the effect that the words are expected to have on the mental state of those who hear (or see) them under certain circumstances."[197]

The closest analogue in the sphere of conduct to these types of harms might be an inappropriate but nonharmful touching of another that amounts to a battery. This type of physical contact could be described as a dignitary injury mediated in part by others' perception of what counts as an acceptable form of touching; in these respects, the perception of touching may be context dependent like speech. However, there is a significant difference between the harms caused by this conduct and by speech. Once society determines which kinds of touching are inappropriate, one can presume that being so touched harms an individual, even if no other viewers witness the event.

Similarly, defamation such as libel, which is not considered protected speech when it meets certain criteria,[198] involves a defendant at least negligently (and often willfully) making false statements that injure a concrete, reputational interest of a plaintiff. Unlike in classic cases of speech harms—although admittedly somewhere in the middle on the spectrum between speech harms and conduct harms—libel involves fault by the defendant that directly affects the material interests and opportunities of a plaintiff without being mediated by a plaintiff's particular, context-dependent emotional response.[199]

---

context dependent and emotionally felt. See id. at 100–01 (describing expected mental effect of harmful speech).

195. See infra section II.B.2 (identifying unique role of speech in helping society change its consideration of what constitutes harm).

196. See, e.g., Am. Booksellers Ass'n v. Hudnut, 771 F.2d 323, 328 (7th Cir. 1985) (invalidating antipornography ordinance justified by concern that "[m]en who see women depicted as subordinate are more likely to treat them so").

197. Schauer, Harm(s) and the First Amendment, supra note 7, at 100–01.

198. See supra notes 95–97 and accompanying text (identifying unique considerations when speech rises to level of libel).

199. See John C.P. Goldberg & Benjamin C. Zipursky, Unrealized Torts, 88 Va. L. Rev. 1625, 1689, 1692 (2002) (explaining despite courts' historical reluctance to redress emotional harm, defamation claims have long been permitted because "there is still a

By analogizing speech harms to conduct harms, courts can protect against the tangible harms caused by speech while preserving speech's specialness. This approach is distinct from but related to regulating the nonspeech components of expressive conduct; it would envision, for example, regulating the burning/destruction component of protestors' burning a draft card even though this also impacts the protest component.[200] The government can, consistent with the First Amendment, control the conduct portion of speech that is classified as expressive conduct so long as the government's target is the conduct and not the actual message.[201] Building on this approach, when free speech consequentialism is unavoidable, courts should account only for the conduct analogues within the speech itself.

This is not to say that all speech contains conduct-like components or that the government can target the harms caused by speech so long as it is not regulating the message. That view would lead to the underprotection of speech that arises from strongly purposivist approaches. Instead, when courts must (and/or do) consider the harms caused by speech at various stages in First Amendment doctrine, the harms that they should specially consider, and the harms that may sometimes outweigh the harm to speech, are the types of harms that are analogous to harms usually caused by conduct.[202]

To be sure, this proposal, which allows for only some harms to matter, acts as a constraint on consequentialism. In some ways, it rejects the fundamental consequentialist premise that outcomes are the only factors relevant to moral evaluations. This proposal can be thought to place a deontological side constraint on free speech consequentialism based on a theory of rights about what courts must protect and what types of harms are not legally cognizable.[203] This side constraint presumes that "speech" has inherent value and also honors deontological rule of law

---

predicate harm in this case—the intangible predicate harm of reputational damage—that provides the foundation for the cause of action").

200. See United States v. O'Brien, 391 U.S. 367, 379–80 (1968) (upholding federal law banning destruction of draft cards because government targeted conduct tampering with selective service procedures, not expressive element of protesting draft).

201. Id. at 376 ("This Court has held that when 'speech' and 'nonspeech' elements are combined in the same course of conduct, a sufficiently important governmental interest in regulating the nonspeech element can justify incidental limitations on First Amendment freedoms.").

202. Cf. Snyder v. Phelps, 131 S. Ct. 1207, 1219 (2011) (refusing to protect individuals from emotionally distressing pure speech on matters of public concern when distress emanated from content of speech). This lack of protection can be extended to all speech, except when the speech at issue produces significant harms that resemble conduct harms.

203. Side constraints normally limit pure consequentialism by prohibiting certain actions/harms even if they lead to the best consequences. See Robert Nozick, Anarchy, State, and Utopia 30–31 (1974) (explaining how side constraint *C* prohibits individuals from doing *C* even in cases where doing *C* would lead to fewer instances of *C* overall). Here, side constraints would limit free speech consequentialism by deeming certain harms irrelevant to the consequentialist calculus.

values associated with upholding the First Amendment. However, there are also consequentialist reasons to adopt this constraint as a rule-based form of consequentialism (establishing a rule that limits case-by-case consideration of consequences, ultimately for consequentialist reasons). Whether speech should be analogized to conduct turns on whether the harms that flow from speech particularly should be treated as special and unregulable. The next section will show why this answer should be yes.

B. *Justifications*

This proposal prevents any free speech consequentialist calculus from treating harms unique to speech the same as harms that flow from conduct. For several reasons, the harms that flow from the expressive aspects of speech should not be weighed against the benefits that flow from speech and are not susceptible to consequentialist analysis. These reasons include the fact that speech harms have special properties that render speech less amenable to regulation; treating speech harms as distinct from conduct harms preserves autonomy; and, as will be discussed first, treating speech harms as distinct from conduct harms honors our collective cultural commitment to the First Amendment.

1. *The First Amendment Establishes a Preference for Speech, Using a Process-Based Definition of Speech.* — The very existence of the First Amendment establishes a preference for speech, defined based on the specialness of speech and perhaps largely independent of its consequences. Both the text of the First Amendment and social and judicial consensus at an abstract level (or at a constitutional "focal point") support treating speech harms as less susceptible to regulation.

Indeed, the First Amendment may eschew free speech consequentialism, especially at the retail level. As the Supreme Court noted in *United States v. Stevens*, "the First Amendment itself reflects a judgment by the American people that the benefits of its restrictions on the Government outweigh the costs. Our Constitution forecloses any attempt to revise that judgment simply on the basis that some speech is not worth it."[204] This proposal is consistent with the First Amendment's preference for speech because it places harms that flow particularly from speech or those that are intertwined with the specialness of speech beyond regulation.

The text of the First Amendment, as compared to other amendments, does not invite harms balancing. Consider the language of the Second Amendment, which provides that "[a] well regulated Militia, being necessary to the security of a free State, the right of the people to keep and bear Arms, shall not be infringed."[205] This Amendment is phrased in consequentialist terms, connecting the right to bear arms to the importance of the security of a free state. The First Amendment, by

204. 559 U.S. 460, 470 (2010).

205. U.S. Const. amend. II.

contrast, provides that "Congress shall make no law . . . abridging the freedom of speech, or of the press."[206] This text also stands in stark contrast to the Fourth Amendment, which prohibits only "unreasonable searches and seizures,"[207] inviting a weighing to determine what is unreasonable.

The text of the First Amendment thus belies the approach of scholars who believe that "the primary goal of modern First Amendment jurisprudence is to . . . identify[] the correct balance between the value of speech and the social harms caused by that speech."[208] Perhaps the First Amendment already incorporates most of its harm balancing at the very highest level of abstraction, with the drafters concluding—either for consequentialist or deontological reasons—that the value of speech generally outweighs its harms, and that undertaking this balancing again and again in particular instances will cause undue damage to speech interests.[209] After all, one of the primary justifications for current First Amendment doctrine is that the government should not decide which speech is high value and which speech is low value; the marketplace of ideas should sort the good speech from the bad.[210]

Of course, the Constitution's text is only one (albeit important)[211] component of constitutionality, especially given the fact that, as established above, First Amendment doctrine is not and should not be absolute. The constraint proposed in this Article also honors society's "focal point" of agreement,[212] interdependent with the common law interpretation of First Amendment's text, that speech is special. While recognizing that free speech consequentialism is unavoidable, this proposal minimizes constant disagreement and costly legislative and

---

206. Id. amend. I. This provision, along with most of the rest of the Bill of Rights, has been incorporated to apply to the states as well. See, e.g., McDonald v. City of Chicago, 130 S. Ct. 3020, 3034 (2010) ("The Court eventually incorporated almost all of the provisions of the Bill of Rights.").

207. U.S. Const. amend. IV.

208. Han, supra note 7, at 1656.

209. See Smolla, supra note 23, at 521 ("The First Amendment itself reflects a judgment by the American people that the benefits of its restrictions on the Government outweigh the costs . . . ." (quoting *Stevens*, 559 U.S. at 470)).

210. See Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting) ("[T]he best test of truth is the power of the thought to get itself accepted in the competition of the market."). Since *Abrams*, "First Amendment doctrine has carried Holmes's laissez-faire marketplace banner more or less faithfully . . . ." Joseph Blocher, Institutions in the Marketplace of Ideas, 57 Duke L.J. 821, 825 (2008).

211. See Curtis A. Bradley & Neil S. Siegel, Constructed Constraint and the Constitutional Text, 64 Duke L.J. 1213, 1227–31 (2015) (describing vaunted status of constitutional text over precedent and custom, and how "[c]onstitutional text seems special" in many ways).

212. See David A. Strauss, Common Law Constitutional Interpretation, 63 U. Chi. L. Rev. 877, 910–15 (1996) (describing how constitutional interpretation follows common law evolution, where Constitution's text establishes focal points of agreement on issues that should remain settled due to costs of constant revisitation).

judicial revisiting of which speech can and cannot be regulated. Reweighing the costs and benefits of speech anew, even with a heavy thumb on the scale toward speech, may eventually degrade our understanding that speech is special. Instead of allowing consequentialism all the way down, with each new instance of speech being weighed and debated anew,[213] courts should accept as a matter of both first principles, and on the basis of administrability concerns, rules that honor America's legal and cultural commitment to the idea that the specialness of speech should be preserved.

Although some scholars argue that free speech "is not a separate value, standing over and above any other in the constitutional hierarchy,"[214] the Constitution should be read so that rights do not conflict. The undesirable alternative is a particularly indeterminate system where, as some scholars propose, the benefits of free speech should be assessed against other constitutional rights, and against concerns as vague and potentially liberty-threatening as "dignity."[215] Despite disagreements at the margins, there is a consensus that speech is special. Expansive protections for free speech are not just a "second-best" solution that allows for the advancement of other values,[216] but a first-best approach that promotes a particular conception of speech as an individual liberty protected against government intrusion. As a result, certain harms—the harms that are inextricably intertwined with the benefits of speech and the harms that we associate with speech—simply do not enter the consequentialist calculus.

Of course, saying speech is special begs the question of what speech is. This Article's approach also has the benefit of defining speech in terms of its distinct characteristics instead of its results (defining speech in terms of which speech is harmful/beneficial). Instead of courts determining substantively which kinds of harms merit governmental intrusion, and when speech actually works against its own instrumental purposes, courts focus on defining what makes speech different from other types of conduct. Courts can then create categories of speech, or determine which harms are unacceptable, based on a coherent definition of "the freedom of speech," instead of deciding cases based on their ideological

---

213. See id. (describing how focal points of agreement breed harmony on issues more controversial and less well settled).

214. Tsesis, supra note 35, at 14.

215. See Neomi Rao, Three Concepts of Dignity in Constitutional Law, 86 Notre Dame L. Rev. 183, 190 (2011) ("Dignity may be appealing as a legal concept precisely because it obscures difficult choices about what we value and the type of freedom and rights we wish to protect.").

216. But see Frederick Schauer, The Second-Best First Amendment, 31 Wm. & Mary L. Rev. 1, 4 (1989) ("The protection of freedom of speech can be seen instrumentally as furthering some background justification, or rationale, or goal.").

preferences and biases, which is often viewed as fundamentally antithetical to free speech values.[217]

Regulating speech based on its properties that can be analogized to conduct harms also allows for courts to adapt to changing technologies and circumstances while preserving expansive free speech protections. Courts should assume, as a baseline, the already accepted categorical exclusions such as obscenity and incitement. This allows precedent to constrain and minimize discord,[218] in accordance with "focal point" theory (additionally, many of the exceptions courts now have can be justified using this Article's approach). Although courts should place a heavy thumb on the scale against creating new categories of unprotected speech, this proposal allows courts to account for harms that are generally not associated with the specialness of speech.

2. *Speech Harms Are Unique in Important Ways.* — Speech harms also should not be compared to harms that generally flow from conduct because speech harms possess two significant qualities: They are likely as easily prevented and alleviated by private citizens as by governmental regulation, and they possess properties intertwined with the benefits of speech.

By now, it has become a veritable cliché that the solution for harms that flow from speech is more speech.[219] And while this cliché is often true,[220] many scholars believe that the marketplace of ideas does not correct errors well.[221] However, there is little indication that lawmakers are better equipped to perceive and correct the harms caused by speech (thus allowing greater scrutiny by the courts than with conduct harms).

---

217. See Lakier, supra note 145, at 2175 (describing how giving judges "opportunity to read their own preferences and ideological commitments into the Constitution" is "threatening to the basic neutrality of First Amendment law").

218. See Shyamkrishna Balganesh & Gideon Parchomovsky, Structure and Value in the Common Law, 163 U. Pa. L. Rev. 1241, 1268–69 (2015) (noting common law precedent undergoes gradual change "emphasiz[ing] . . . restraint"); Tsesis, supra note 35, at 4–5 (noting ability of precedent to constrain, even when it does not match ideals about how law should operate, is important force allowing law to adapt to change while preserving stability).

219. Martin H. Redish, Tobacco Advertising and The First Amendment, 81 Iowa L. Rev. 589, 606 (1996) ("At least since the famed concurring opinion of Justice Brandeis in *Whitney v. California*, it has been well accepted that the answer to supposedly harmful speech is not governmental suppression, but rather more speech.").

220. As one example, consider the success of antitobacco campaigns, statistically correlated with a reduction in smoking. See, e.g., Marc Kaufman, Anti-Smoking Campaign Shows Dramatic Results, Wash. Post (Mar. 1, 2000), http://www.washingtonpost.com/archive/politics/2000/03/01/anti-smoking-campaign-shows-dramatic-results/f536fcbb-d9d5-4c79-8277-e30d1bcab834/ [http://perma.cc/2MX5-F69R] (chronicling antitobacco campaign).

221. See, e.g., Adrian Vermeule, The Invisible Hand in Legal and Political Theory, 96 Va. L. Rev. 1417, 1447 (2010) (describing criticism that marketplace of ideas produces truth).

Even Brian Leiter, who doubts the specialness of most speech,[222] acknowledges that the alternative—letting the government decide when speech harms are too great—raises the problem of the epistemic arbiter.[223] The government may be no more capable of sorting out harmful speech from helpful speech, and mitigating the harms of speech, than individuals are. Governmental regulation of the expressive elements of speech renders speech vulnerable to censorship based on impermissible governmental motives, like viewpoint discrimination or arbitrary and subjective understandings of harm, and creates more difficult line-drawing problems.[224]

State intervention to remedy speech harms is also problematic because the question of whether a particular instance of speech causes harm changes with time, and with society's perception of that speech.[225] Because speech is highly context dependent,[226] the harm caused by speech is more malleable than the harm caused by conduct. Indeed, outlawing speech as being harmful often compounds the sense of harm listeners feel from the speech;[227] this phenomenon argues in favor of the strong-listener model.[228] It is far more difficult, and often counter-productive, to weigh the harms caused by speech and balance them against speech's virtues.

More critically, the very harms that flow from speech are intertwined with the way we resist speech-based harms. The malleability of speech harms and the diffuse nature of those responsible for speech harms allow speech to serve unique and important functions in society. Speech plays a special role in catalyzing action and change through diffuse parties without itself manifesting change. Speech is how we convince people that

---

222. See Leiter, Case Against Free Speech, supra note 78, at 3 ("In a slogan: most non-mundane speech people engage in is largely worthless, and the world would be better off were it not expressed.").

223. See id. at 34 ("There is only one serious argument against regulation of speech[:] . . . [that] we do not have a reliable epistemic arbiter, and, moreover, any attempt to designate one runs the risk of sacrificing all the other goods associated with free speech insofar as the arbiter is unreliable.").

224. After all, each individual has his own idea about which speech is particularly harmful.

225. See supra notes 192–195 and accompanying text (explaining difficulty in regulating speech harms due to changing perception of harms).

226. See Kathleen M. Sullivan, Resurrecting Free Speech, 63 Fordham L. Rev. 971, 983 (1995) [hereinafter Sullivan, Resurrecting Free Speech] ("[S]peech depends on language, which is a collective social creation over time. For another, the concepts and ideas that we deploy likewise come from a collective culture, an accretion of past lives in which we are steeped and educated, whether we like it or not.").

227. Listeners will not only experience the emotional upset from the speech itself, but will feel indignant that their legal rights have been violated. See, e.g., infra notes 321–330 and accompanying text (discussing trigger warnings in universities and academic freedom rights).

228. See supra section II.A (discussing harms to listeners caused by speech and arguing for courts to analogize conduct harms to speech harms).

the harms that flow from speech are harmful, and also how we change others' (and our own) perception of what is harm and what it means to be harmed. Speech is one of the few ways society evolves in its consideration of what constitutes harm.[229]

Conduct harms, by contrast, often are direct, immediate, and caused by the party engaging in the conduct, and elicit physical or tangible responses or impairment of immediate, material interests. Laws restricting conduct can thus target the illegal change, or action causing damage, while allowing the underlying speech to flourish and perhaps foster progress in the future. This partially explains why speech depicting illegal activity is legal even if the underlying conduct can be regulated.[230] Even speech that most people find loathsome or incendiary affects social norms and attitudes without directly targeting tangible interests.

Conduct (or even speech) that is regulated via criminal, tort, or agency law, or even discouraged through social stigma, is often avoided,[231] and society generally accepts that this conduct is harmful. Through the medium of speech, we can espouse views that particular behavior is not harmful without needing to engage in that conduct, and without suffering the penalties, to illustrate the point.[232] Of course, there may be social stigma also associated with advocating in favor of controversial conduct, but the First Amendment protects the right to both associate with others to amplify controversial speech and to speak anonymously.[233] Free speech thus allows society's consequentialist calculations about harm to evolve and progress;[234] this mechanism would thus work best if it were largely insulated from the consequentialist calculus.

---

229. See Sullivan, Resurrecting Free Speech, supra note 226, at 984 ("[L]aws . . . provisionally bind our conduct . . . . [A]s long as the speech/conduct distinction holds, we are still free in the meantime to vent, and rage, and scream, and yell, or even politely argue that the laws that bind our conduct are misconceived and ought best be changed.").

230. See United States v. Stevens, 559 U.S. 460, 482 (2010) (holding depictions of animal cruelty are protected by First Amendment).

231. See Cass R. Sunstein, Social Norms and Social Roles, 96 Colum. L. Rev. 903, 953 (1996) (identifying "function of law in expressing social values and in encouraging social norms to move in particular directions"); see also Mark A. Edwards, Law and the Parameters of Acceptable Social Deviance, 97 J. Crim. L. & Criminology 49, 50–51 (2006) (arguing norms of social behavior sometimes have greater influence on behavior and discretionary enforcement decisions than does law).

232. See Sullivan, Resurrecting Free Speech, supra note 226, at 984 ("[T]hrough speech, we also forge the strands of opposition to any set of decisions that we have provisionally embraced to regulate our conduct through the processes of government and the enactment of law.").

233. See McIntyre v. Ohio Elections Comm'n, 514 U.S. 334, 342 (1995) ("[A]n author's decision to remain anonymous, like other decisions concerning omissions or additions to the content of a publication, is an aspect of the freedom of speech protected by the First Amendment.").

234. See Sullivan, Resurrecting Free Speech, supra note 226, at 984 ("[I]t is critical to our autonomy individually, or to our self-government as a whole, that we be able always to say, 'Maybe we ought to change our conduct regulations.' . . . [S]peech . . . [is] the only

Because certain harms caused by speech change with time and context, and, in fact, change as society perceives harms differently, factoring these types of harms into the consequentialist analysis will cause courts to place too much weight on harms that are unstable or diminishing. When speech triggers liability precisely because it is emotionally distressing, this distress is often based on the unpleasant message or views of the speaker.[235] In addition to predicating liability on viewpoint, allowing emotional harms caused by speech to determine regulability undermines a primary purpose of speech (i.e., to alter emotions, and to elicit responses that may be negative or distressing).

3. *Autonomy and Responsibility.* — Distinguishing speech harms from conduct harms also promotes the view that adults generally have the capacity to lead autonomous lives and, absent extraordinary influences, they in fact do so. Autonomy, or at least autonomy in the "negative" sense that would prevent governmental suppression of speech,[236] requires the normative conception of people as largely capable of resisting coercion, emotional manipulation, or temporary distortion of judgment that stem just from speech. A social climate that accepts individuals as autonomous in this way is more likely to both assign them responsibility for their actions and allow them agency to make voluntary decisions. A climate in which harms that flow from speech are regulable denies agency in ways that regulating conduct harms does not, because it presumes that others can overtake the sphere of our own minds, and it absolves us of responsibility for managing our emotions.[237] This climate has negative implications for our abilities to consent to voluntary arrangements and also undermines ideas of legal culpability in both tort and criminal law.

In a pro-agency, pro-autonomy world, it is unfair and unnecessarily restrictive to hold a speaker responsible for acts neither intended by nor condoned by the speaker, because speech harms are often caused by diffuse parties or intermediaries.[238] Speech has unique characteristics that

---

legitimate medium we recognize—by which we advance from one regime of conduct regulation to another.").

235. See Snyder v. Phelps, 131 S. Ct. 1207, 1219 (2011) (overturning $5 million jury award for IIED, based on hateful, antigay protests at military funeral, where 8-1 majority concluded "any distress occasioned by Westboro's picketing turned on the content and viewpoint of the message conveyed").

236. See Richard H. Fallon, Jr., Two Senses of Autonomy, 46 Stan. L. Rev. 875, 881 (1994) ("Describing autonomy as freedom from coercion, manipulation, or temporary distortion of judgment, negative libertarians characteristically claim that respect for autonomy requires very broad freedoms of speech.").

237. See Goldberg, supra note 31, at 837–41 (demonstrating why "[e]xpansive emotional duties to others' emotional well-being would also reduce autonomy for both the putative plaintiff and the would-be defendant in all manner of consensual relationships, from employer–employee to friends").

238. This leaves open, for example, the possibility of holding speakers liable for how-to guides involving building bombs or committing acts of violence, but only if the speaker directed or intended a target for the how-to guide.

create important reasons not to place blame on the speaker for harm that arises. Tort law, which already contains sophisticated frameworks for conceptualizing blame and responsibility, provides a useful concept: proximate cause. Tort law requires that a defendant be both a but-for cause and a proximate cause (or culpable party) in order to impose liability.[239] This sense of culpability is often based on what is reasonably foreseeable, in a normative sense, and contingent upon the foreseeability of the actions of other parties who contribute to the harm.[240]

Foreseeability concepts break down in the face of speech harms. When the harm caused by speech is context dependent, it is more difficult to foresee the harm it will cause, as that harm changes over time and in different contexts. And, conversely, it is usually predictable that someone, even a reasonable person, may be bothered by certain speech, even if simply because he objects to the speech. Instead of blaming the speaker, when the harm caused by speech is largely emotional in nature, the listener should be deemed responsible for managing his own emotional response and thus mitigating the harm caused.[241] When the harm is perpetuated by diffuse agents, for example in a context of reputational harm where the harm is created when each new audience member is exposed to the speech, there are intervening causal forces that militate against holding the speaker liable except in cases of malicious falsity. Diffusion also happens when harm is perpetuated in a way wholly separate from the speaker's harm, for example when a viewer of a violent movie copies an act committed in that movie.

Analogizing speech harms to conduct harms would preserve the specialness of speech and its important benefits to society. Regulating speech because it agitates others through diffuse processes, or because it causes emotional upset to a listener, necessarily undermines the reason that speech is special—because of its ability to impact others in an attenuated way, through the cognitive or emotional processes of a listener or through diffuse parties ruminating upon the speech, without causing specific or tangible harms. However, when speech begins to resemble conduct, such as when it impairs discrete, material interests through direct processes and through the fault mostly of the speaker, then courts should consider those conduct-like harms in their consequentialist calculus.

---

239. See Jill E. Fisch, Cause for Concern: Causation and Federal Securities Fraud, 94 Iowa L. Rev. 811, 830–31 (2009) (describing differences between factual and legal causation).

240. See, e.g., Edwards ex rel. Fryover v. Anderson Eng'g, Inc., 251 P.3d 660, 665 (Kan. Ct. App. 2011) (holding "intervening acts or consequences cannot be said to be foreseeable" when "negligent acts of others . . . broke the connection between the initial negligent acts and the harm caused").

241. See Goldberg, supra note 31, at 816 (explaining and justifying disparate treatment in tort law of physical and emotional injury based on duty to reasonably regulate one's own emotional well-being).

C. *Addressing Objections to Free Speech Consequentialism*

The approach of preserving strong, aspirationally neutral free speech protections by analogizing speech harms to conduct harms raises several concerns. Three arguments merit special attention: first, many instances of speech can likely be analogized to conduct; second, emotional harms are significant and therefore should be accounted for in free speech consequentialism; third, not accounting for speech-specific harms has unfortunate distributional consequences.

First, although it is feasible to analogize all speech harms to conduct harms, some analogies will be more apt than others. To avoid having the solution to unprincipled free speech consequentialism contain just as many pitfalls for subjectivity and bias, courts should analogize speech harms to conduct harms only in rare cases where the speech harm closely incorporates a harm usually caused by conduct, or where there is a tangible and compelling interest that can be tied to the harm from speech, like a reputational interest or an interest in education.

Although scholars often compare the emotional distress caused by speech to conduct, this is an improper analogy. A "verbal assault" contains very few similarities to a physical battery, except for the fact that both cause pain, which may be sharp at first and diffuse over time. Although there is a connection between the mind and the body, and the pain caused by words may ultimately manifest as physical symptoms, the harm to emotional tranquility is conceptually different from the harm to bodily integrity.[242] This difference between physical and emotional pain, in fact, underlies the speech/conduct distinction,[243] without which, either all speech or no conduct could be regulable.

Of course, an objection can be raised that even the harms characteristic of speech—like context-dependent harms, emotional harms, or harms caused by diffuse parties—should be incorporated into free speech consequentialism. In particular, many believe that courts should intervene more to protect people's emotional well-being.[244] After all, one could make a reductivist argument that the only reason we care about any possible interest, be it bodily, reputational, property, or privacy, is because ultimately we care about how that interest impacts our emotional well-being, or overall happiness.

On some level, this may be true, notwithstanding the idea that individuals may care about particular interests, ideals, or values unrelated to how they impact their personal happiness. Even if true, many aspects of the law, fault, and responsibility depend on the notion that we are the

242. See id. at 823 (discussing courts' conception of physical injuries "as more worthy of protection" than "stand-alone emotional claims").

243. See Sullivan, Resurrecting Free Speech, supra note 226, at 976 (drawing distinction between mind and body as similar to that of speech and conduct).

244. See supra note 36 and accompanying text (presenting view Court treats harms caused by speech dismissively without genuinely addressing evidence).

intermediaries between other interests and our nebulous sense of happiness.[245] The law should protect tangible, identifiable interests, but cannot ultimately make us happy. John Goldberg and Benjamin Zipursky's article *Unrealized Torts* eloquently describes this approach as it applies to the tort of false imprisonment:

> Consider, for example, the tort of false imprisonment. That tort involves intentionally creating a situation in which a plaintiff would reasonably take herself to be confined. . . . Of course, the primary effect this tort has on an ordinary person is to cause them to experience fright, anger, embarrassment, frustration, and/or other negative feelings. But a cause of action for false imprisonment does not identify these as the injuries, nor does it identify the causing of them as the wrong. Instead, false imprisonment is a dramatic example of depriving someone of an important liberty, the liberty of movement. The essence of the injury is the deprivation of that liberty.[246]

Tying the harms from speech to identifiable interests also ensures that the harms are of a magnitude and kind that society is abstractly prepared to prevent. As examples, in *United States v. Alvarez*, Justice Kennedy's plurality opinion credited as a governmental interest the dilution of "the public's general perception of military awards," but did not incorporate the "anger and frustration" that may be experienced by actual recipients of military awards.[247] In the context of endorsing civil orders against harassment, as another example, one scholar argues that civil harassment statutes should protect not just against a general sense of unease, but should be explicitly tied to unconsented contact or surveillance that threatens privacy.[248] He cites, as too broad and untethered to specific interests, a Washington statute that reads, "'Unlawful harassment' means a knowing and willful course of conduct directed at a specific person which seriously alarms, annoys, harasses, or is detrimental to such person, and which serves no legitimate or lawful purpose."[249] Explicitly using terms like

---

245. See Goldberg, supra note 31, at 857–58 ("The physical/emotional distinction, and a modern justification based on duties to maintain our own emotional health, should consider defendants less responsible, and therefore less at fault, for others' emotional distress.").

246. Goldberg & Zipursky, supra note 199, at 1692.

247. 132 S. Ct. 2537, 2549 (2012) (plurality opinion) (Kennedy, J.).

248. See Aaron H. Caplan, Free Speech and Civil Harassment Orders, 64 Hastings L.J. 781, 826–27 (2013) ("The key to a constitutionally acceptable standard for civil harassment is the concept of unconsented contact or surveillance that threatens safety and privacy.").

249. Id. at 827 (quoting Wash. Rev. Code § 10.14.020(1) (2012)). Harassment should also require reasonableness and pervasiveness components, see infra notes 284–286 and accompanying text (discussing judicial treatment of harassment claims relating to speech issues on college campuses), but Caplan recognizes that "[s]erious constitutional questions are sure to arise under the California model, which authorizes injunctions (against anything) in response to (unspecified) behavior that makes others feel (generally) bad." Caplan, supra note 248, at 826.

unconsented contact and repeated invasions of privacy, thus "specifying the emotional harms that come from loss of safety and privacy—as opposed to sketchily defined emotional distress—ensures that a petitioner's reasons for avoiding contact are those that society is prepared to endorse."[250]

Further, everyone has a different sense of which emotional responses are justified and which are disproportionate. Deciding whose emotional responses are valid enough that speech should be suppressed really would turn First Amendment law into crass politics. Accounting for the emotional harms caused by speech would undermine one of the primary purposes of speech: to arouse emotions, to stir people to thought and anger, and ultimately, to perhaps stir them to action.[251]

This does not mean that, as a society, we should undermine others' emotional experience. As society becomes increasingly attuned to the importance of emotional and mental health, we should recognize emotional injury as significant and serious. However, although each individual should be entitled to his emotional response, that response should not dictate the First Amendment rights of others. Emotions need not be pathologized,[252] but we might all be better off accepting that they cannot be the sole source of legal rights, especially when they conflict with free speech protections.[253]

Unfortunately, as many scholars note, from a distributional perspective, the emotional burdens of speech may not be shared equally. This is a significant concern, especially when the burdens of speech fall on particularly vulnerable members of the population, or those who have been historically oppressed. Some feminist scholars argue that anonymous speech has made the Internet a hostile place, especially for women.[254] Proponents of laws regulating hate speech note that this

---

250. Caplan, supra note 248, at 827. Of course, privacy, when based on its emotional effects without being associated with a tangible interest, is almost as subjective an interest as emotional tranquility and should be invoked sparingly. See infra notes 293–294 (discussing challenges associated with using privacy claims to police revenge porn).

251. See Cohen v. California, 403 U.S. 15, 25–26 (1971) ("[M]uch linguistic expression serves a dual communicative function: it conveys not only ideas capable of relatively precise, detached explication, but otherwise inexpressible emotions as well. In fact, words are often chosen as much for their emotive as their cognitive force.").

252. See Julie Holland, Opinion, Medicating Women's Feelings, N.Y. Times (Feb. 28, 2015), http://www.nytimes.com/2015/03/01/opinion/sunday/medicating-womens-feelings.html (on file with the *Columbia Law Review*) ("Women's emotionality is a sign of health, not disease; it is a source of power.").

253. See infra Part III (examining problems posed by basing legal rights solely on emotional response in context of hate speech, revenge porn, and trigger warnings).

254. See Danielle Keats Citron, Law's Expressive Value in Combatting Cyber Gender Harassment, 108 Mich. L. Rev. 373, 396–97 (2009) ("[O]nline abuse inflicts significant economic, emotional, and physical harm on women in much the same way that workplace sexual harassment does."); see also Conor Friedersdorf, When Misogynistic Trolls Make Journalism Miserable for Women, Atlantic (Jan. 7, 2014), http://www.theatlantic.com/politics/archive/2014/01/when-misogynist-trolls-make-journalism-miserable-for-women/

speech is particularly corrosive to the psychological well-being of racial minorities, whose ability to engage in counterspeech may consequently be diminished.[255] Jeremy Waldron argues that speech that disparages individuals or groups on the basis of race, gender, religion, sexual orientation, or other status-based characteristics undermines a sense of equal citizenship for the targets of the speech.[256]

These concerns, while powerful, are ultimately not ones that free speech law should countenance. These distributional arguments often require resorting to the most easily manipulated and potentially subversive form of free speech consequentialism—the view that certain speech should be suppressed so that more speech can thrive overall, or so that certain people can speak more freely (i.e., we should suppress speech to promote speech). This argument is generally based on the premise that certain speech causes some to feel so uncomfortable, emotionally distressed, or alienated that they will self-censor, not because of government action, but because of emotional upset caused by a private party. The view that some speech should be suppressed so other speech can thrive threatens our free speech regime because it depends on difficult empirical and normative judgments about how suppression of speech impacts other speech and more broadly, how suppression of speech impacts the welfare of members of particularly vulnerable groups of society.

It is unclear whether members of minority groups fare better in societies where speech damaging to these groups is aired or in cultures that prohibit this type of speech and force it underground. There is reason to believe that suppressing the types of speech most damaging to women and minorities may not actually benefit members of these groups in the long term. Some feminists, generally associated with third-wave feminism, convincingly argue that the association of women with emotional trauma, and the law's efforts to specifically cater to women's emotional distress, has actually deprived women of autonomy and the ability to make their own choices.[257] Further, instead of feeling victimized, many are empowered by collective efforts to speak out against expressions of bigotry.[258] The U.S. approach of tolerating the intolerant may actually

---

282862/ [http://perma.cc/48LZ-MCKN] (describing chilling effect of online misogyny on female bloggers and journalists).

255. See, e.g., Mari J. Matsuda, Public Response to Racist Speech: Considering the Victim's Story, 87 Mich. L. Rev. 2320, 2357–58 (1989) (identifying characteristics defining racist hate speech for which counterspeech is insufficient).

256. See Jeremy Waldron, 2009 Oliver Wendell Holmes Lectures, Dignity and Defamation: The Visibility of Hate, 123 Harv. L. Rev. 1596, 1609–12 (2010) (conceptualizing hate speech as assault on "dignity," defined as individual social status).

257. See Jeannie Suk, The Trajectory of Trauma: Bodies and Minds of Abortion Discourse, 110 Colum. L. Rev. 1193, 1197, 1203–06, 1214–15 (2010) (arguing association between women and emotional trauma, advanced by feminist movement of 1970s, may be responsible for laws that diminish women's agency, like partial-birth abortion ban).

258. In March of 2011, a student at UCLA posted a racist "rant" about the behavior of Asian students at her school on YouTube, and this unintentionally viral video inspired a

make society overall more tolerant and more welcoming to marginalized groups.[259]

If scholars who wish to suppress some speech to promote the speech of others are wrong about how speech actually affects the most vulnerable members of society, speech is being suppressed regardless. This unflinchingly consequentialist approach would render speech forever vulnerable to the claim that the speech of some threatens the speech of others, and that the speech of some is more valuable than the speech of the others. Regulations targeting speech damaging to particularly vulnerable groups are also, unfortunately, easily manipulated to serve as proxies for viewpoint discrimination.[260] This does not mean that vulnerable and historically oppressed groups must suffer the extra emotional burdens from speech in silence. The government can speak for itself in denouncing racism and sexism, so long as it allows private citizens to express their own views.[261] This would also allow the government to help

---

torrent of creative, and often humorous, parodies challenging stereotypes of Asians, and even became the name of a Chinese take-out service. See Alexandra Wallace's 'Ching Chong Ling Long' Rant Inspires Takeout Service, Huffington Post Los Angeles (Aug. 21, 2011, 5:12 AM), http://www.huffingtonpost.com/2011/06/21/alexandra-wallaces-ching-_n_881381.html [http://perma.cc/2TRN-WXVM] (describing video and takeout service).

259. France's attempt to ameliorate intolerance on the basis of religion serves as a contrasting example. Laws in France aimed at prohibiting religious discrimination in schools and promoting strong separation between church and state banned individual pupils from wearing "ostensible religious symbols or clothing in public schools." Ofrit Liviatan, From Abortion to Islam: The Changing Function of Law in Europe's Cultural Debates, 36 Fordham Int'l L.J. 93, 97 (2013). This law, designed to ameliorate alienation and intolerance and in accordance with France's policy of secularism, actually perpetuated these problems in the view of many. See John R. Bowen, Why The French Don't Like Headscarves 1 (2007) ("Although worded in a religion-neutral way, everyone understood the law to be aimed at keeping Muslim girls from wearing headscarves in school.").

260. To list two noteworthy examples, both taking place at universities, a student-employee at Indiana University-Purdue University Indianapolis (IUPUI) was found guilty of racial harassment after a coworker was offended by his reading a book entitled "Notre Dame vs. the Klan: How the Fighting Irish Defeated the Ku Klux Klan" during his work breaks. See Deanna Martin, IUPUI Says Sorry to Janitor Scolded over KKK Book, Huffington Post (July 14, 2008), http://www.huffingtonpost.com/huff-wires/20080714/kkk-book-apology/?m=true [http://perma.cc/R45T-SJ5N]. An administrator at IUPUI wrote a letter to the student-employee claiming, "[y]ou used extremely poor judgment by insisting on openly reading the book related to a historically and racially abhorrent subject in the presence of your black co-workers." Id. The book was actually a historical account celebrating the defeat of the Klan. Id. In the second incident, Bucknell University shut down a student group's "affirmative action bake sale," a form of protest popular among conservative students who oppose racial preferences in college admissions. See Ashby Jones, Bucknell and the 'Affirmative Action Bakesale', Law Blog, Wall St. J.: Law Blog (June 23, 2009), http://blogs.wsj.com/law/2009/06/23/bucknell-and-the-affirmative-action-bakesale/ (on file with the *Columbia Law Review*). The university claimed that the symbolic gesture of the bake sale's pricing scheme being tied to membership in a particular racial group was itself discriminatory because members of certain minority groups were charged less for the baked goods. Id.

261. See Nadine Strossen, Regulating Racist Speech on Campus: A Modest Proposal?, 1990 Duke L.J. 484, 565 (recognizing equality values may be promoted most effectively through nonregulation of hate speech).

create a more inclusive society and mitigate Waldron's concern that hateful, bigoted speech robs individuals of "basic social standing, of the basis of their recognition as social equals, and of their status as bearers of human rights and constitutional entitlements."[262] A sharp divide between government as speaker and individual as speaker allows the government to dissociate itself from bigoted speech and project its commitment to equal rights for all citizens.[263] Actual threats of imminent bodily harm are also regulable, as is speech intended to intimidate.[264] And the social stigma that attaches to bigoted speech becomes stronger and stronger over time.[265] This means that not only will there be less bigoted speech, and more counterspeech, but the emotional harm caused by this speech diminishes as the intolerant become more and more socially marginalized.

Just because, in rare cases, courts should account for the harms caused by speech that can be analogized to conduct, does not mean that these harms will generally outweigh the benefits of speech. Indeed, a robust, principled First Amendment jurisprudence depends on the strictness of strict scrutiny. The final section of this paper examines several timely free speech issues through the lens established above.

III. APPLICATION TO CURRENT FREE SPEECH PROBLEMS

Although the Supreme Court appears committed to protecting First Amendment freedoms,[266] our culture is changing. Within and apart from legal decisions and discussions, there is a greater focus on mental health and protecting individuals' emotional needs.[267] Increased attention to this

---

262. Waldron, supra note 256, at 1610.

263. See Pleasant Grove City v. Summum, 555 U.S. 460, 467–68 (2009) ("A government entity has the right to speak for itself. [I]t is entitled to say what it wishes, and to select the views that it wants to express." (alteration in original) (citations omitted) (internal quotation marks omitted) (first quoting Bd. of Regents of Univ. of Wis. Sys. v. Southworth, 529 U.S. 217, 229 (2000); and then quoting Rosenberger v. Rector of Univ. of Va., 515 U.S. 819, 833 (1995))).

264. See R.A.V. v. City of St. Paul, 505 U.S. 377, 391, 396 (1992) (invalidating St. Paul ordinance applying only to fighting words "that insult, or provoke violence, 'on the basis of race, color, creed, religion or gender'").

265. At some point, social stigma becomes so strong that society may actually be suppressing speech, although not in a way that is nearly as coercive as government censorship, and thus not involving the same sort of intrusion on liberty. So long as anonymous speech is protected and associational rights are maintained, there is also a degree of protection for unpopular views against widespread social condemnation.

266. See, e.g., United States v. Alvarez, 132 S. Ct. 2537, 2551 (2012) (Kennedy, J.) (plurality opinion) (holding Stolen Valor Act infringes on free speech); Snyder v. Phelps, 131 S. Ct. 1207, 1219–20 (2011) (holding First Amendment barred petitioner recovery for IIED or intrusion upon exposure to antigay funeral picketing); United States v. Stevens, 559 U.S. 460, 482 (2010) (holding criminalization of possession of depictions of animal cruelty substantially overbroad under First Amendment).

267. See Goldberg, supra note 31, at 814 & n.19 ("Scientists' and scholars' objections to the physical/emotional distinction are gaining momentum as our nation is becoming increasingly concerned with mental health.").

important aspect of health and well-being risks as an unfortunate byproduct the view that avoiding emotional discomfort should trump robust discussion, the expansion of knowledge, and individual freedoms. Instead of judging behavior primarily by the feelings the behavior engenders, there should be ways to determine what is "reasonable," and what rights we have as individuals. When analyzing free speech issues, both as part of legal doctrine and when considering what this Article terms private "free speech values,"[268] analogizing speech harms to conduct harms can help reset this balance. Although the First Amendment prevents the government from restricting speech in order to precipitate cultural shifts, preserving strong, robust free speech jurisprudence against government imposition may have side benefits. This includes safeguarding against the erosion of a strong cultural commitment to free speech values, even in private interactions.

This Part addresses important and emerging issues in free speech doctrine and culture, such as campus hate speech, trigger warnings, and the regulation of dangerous speech. The methods proposed above can begin to resolve several controversial questions implicating the role of the courts in accounting for harms of verbal assault, dissemination-based harms, participant harms, and harms of advocacy.

A. *Blending the Approaches: The Actual Treatment of Campus Hate Speech*

Although courts should avoid free speech consequentialism, there will be exceptions. There may also be places where the harm from speech cannot be ignored. Consider, for example, students at a public university, who are constantly exposed to speech that is constitutionally protected but perhaps highly harmful, like a swastika.[269] State universities, as public fora,[270] cannot, consistent with current First Amendment jurisprudence, place a blanket prohibition on racist symbols on campus.[271] Perhaps uni-

---

268. The term "free speech values" (as compared to "free speech rights" or "free speech doctrine") refers to the rationales and justifications behind the First Amendment that exist independent of the government's role in policing speech. For example, the idea that people should feel comfortable expressing controversial views, and the idea that we should all keep open minds and use discussions to search for truth, are free speech values.

269. See Nat'l Socialist Party of Am. v. Skokie, 432 U.S. 43, 43–44 (1977) (holding First Amendment safeguards are necessary in case where state court issued injunction prohibiting "(m)arching, walking or parading in the uniform of the National Socialist Party of America; (m)arching, walking or parading or otherwise displaying the swastika on or off their person").

270. See Widmar v. Vincent, 454 U.S. 263, 267 & n.5, 274–77 (1981) (noting universities possess many characteristics of public forums).

271. See McCauley v. Univ. of the Virgin Is., 618 F.3d 232, 248 (3d Cir. 2010) (holding school regulation banning "offensive" signs at sporting events was overbroad); see also Saxe v. State Coll. Area Sch. Dist., 240 F.3d 200, 206 (3d Cir. 2001) ("[T]he free speech clause protects a wide variety of speech that listeners may consider deeply offensive, including statements that impugn another's race or national origin or that denigrate religious beliefs.").

versities could regulate the use of swastikas in dormitories, because the government owns the property and is serving a role as landlord more than regulator. However, even dormitories may be considered at least a limited public forum, so viewpoint-based discrimination is impermissible,[272] and the rest of campus is considered a public forum.[273] What if a significant number of students constantly walked around university property holding up swastikas, not in a crowd or in a parade that was disruptive, such that a free speech exception might apply, but just as a matter of course? What if students were exposed to swastikas every time they traversed campus to attend class?

Courts that find this situation intolerable have several options. They can create an entirely new, unprotected category of speech at a high level of abstraction, perhaps called "hate speech."[274] Or courts can conduct Justice Breyer-like weighing to determine that, because of the costs and benefits of this type of speech, the speech at issue is subject only to intermediate scrutiny, and then hold that any campus regulation targeting this speech survives intermediate scrutiny. Or courts can continue to believe hateful speech is protected and apply strict scrutiny to constraints on this speech and hold that the university's regulation survives strict scrutiny.

There are costs and benefits to each approach. Holding that a regulation survives strict scrutiny has the potential to dilute the strict scrutiny standard, potentially undermining future free speech protections. The application of strict scrutiny happens at such a low level of abstraction that a high standard is necessary to avoid the subjective, political vices of free speech consequentialism.

That said, courts should avoid creating new, unprotected categories of speech at the wholesale level, because this will allow a wide swath of speech to be regulated based on content. Indeed, the category of "hate

---

272. See DeJohn v. Temple Univ., 537 F.3d 301, 314 (3d Cir. 2008) ("Because overbroad harassment policies can suppress or even chill core protected speech, and are susceptible to selective application amounting to content-based or viewpoint discrimination, the overbreadth doctrine may be invoked in student free speech cases.").

273. See Justice For All v. Faulkner, 410 F.3d 760, 767–69 (5th Cir. 2005) ("The Institutional Rules, like the regulations in *Hays County Guardian*, clearly evince an intent to maintain the Austin campus as a designated forum for student expression, subject only to time, place, and manner regulations and a small number of enumerated content-based restrictions." (citing Hays Cty. Guardian v. Supple, 969 F.2d 111 (5th Cir. 1992))).

274. Many scholars have proposed this idea, but their arguments have ultimately been unavailing. See, e.g., Matsuda, supra note 255, at 2356–61 (proposing framework to exclude racist speech from First Amendment protection); see also David Goldberger, Sources of Judicial Reluctance to Use Psychic Harm as a Basis for Suppressing Racist, Sexist and Ethnically Offensive Speech, 56 Brook. L. Rev. 1165, 1166 (1991) ("One of the most puzzling aspects of the judiciary's stance protecting racially, ethnically and sexually offensive public speech is that it has . . . often conceded that offensive speech causes emotional harm while ruling that such speech, nonetheless, is entitled to constitutional protection.").

speech" may be based on an aversion to particular viewpoints.[275] Speech that could be labeled "hate speech," unlike in every other Western democracy, is a protected category of speech.[276] This category has been deployed to suppress unpopular views in other countries, where hate speech is unprotected.[277] Courts in the United States will not allow the government, for example, to arrest people for satirically praising an act of terrorism in order to criticize the government, as France does.[278] Because of slippery slope concerns, and due to fear of impermissible governmental viewpoint discrimination, creating a category of hate speech would erode too much of our robust free speech protections.

Categorically declaring hate speech unprotected at the wholesale level also has the potential to create slippery slope problems for other entire categories of speech. This was the Court's fear in *United States v. Stevens*, where during oral argument several Justices mused that if depictions of actual cruelty to animals were considered an unprotected category of speech, then Congress should also be able to regulate other depictions of illegal activity, like fights to the death that are legal in other countries,[279] depictions of ethnic cleansing in other countries,[280] "the human sacrifice channel,"[281] or even depictions of Adolph Hitler.[282]

---

275. See DeAngelis v. El Paso Mun. Police Officers Ass'n, 51 F.3d 591, 596–97 (5th Cir. 1995) (noting antidiscrimination laws "applied to . . . harassment claims founded solely on verbal insults, pictorial or literary matter, . . . impose[] content-based, viewpoint-discriminatory restrictions on speech"); see also *Saxe*, 240 F.3d at 204 (noting "[t]here is no categorical 'harassment exception' to the First Amendment's free speech clause").

276. See Michel Rosenfeld, Hate Speech in Constitutional Jurisprudence: A Comparative Analysis, 24 Cardozo L. Rev. 1523, 1541–54 (2003) (comparing Canadian, British, and German approaches to hate speech in constitutional jurisprudence).

277. See Jonathan Rauch, Kindly Inquisitors 1–2 (1993) (recounting examples from France, Australia, Austria, and Canada); Adam Liptak, Hate Speech or Free Speech? What Much of West Bans Is Protected in the U.S., N.Y. Times (June 11, 2008), http://www.nytimes.com/2008/06/11/world/americas/11iht-hate.4.13645369.html (on file with the *Columbia Law Review*) (discussing Canadian magazine's hate speech case).

278. See Krishnadev Calamur, Controversial French Comedian Arrested over Facebook Post on Paris Attacks, NPR: The Two-Way (Jan. 14, 2015, 11:46 AM), http://www.npr.org/blogs/thetwo-way/2015/01/14/377201227/controversial-french-comedian-arrested-over-facebook-post-on-paris-attacks [http://perma.cc/GYH8-DZ86] (discussing French arrest of comedian who praised Jewish grocery store shooting).

279. For example, Justice Alito posited:

> Suppose that I am an aficionado of the sort [of] gladiatorial contests that used to take place in ancient Rome, and suppose that some—Rome or some other place decides that it wants to make money by staging these things and selling videos of them or broadcasting them live around the world. Do you have any doubt that that could be prohibited?

Transcript of Oral Argument at 23, United States v. Stevens, 559 U.S. 460 (2009) (No. 08-769).

280. Id. at 48 (reporting Justice Alito's hypothetical).

281. Id. at 46 (same).

282. Id. at 47 (reporting Justice Scalia's hypothetical).

Perhaps the most corrosive option, however, would be the doubling down on free speech consequentialism, like the concurrence in *Alvarez*, which possesses the vices of both of the other approaches.[283] Here, a court would have to first decide to create a pseudo category of speech, hateful symbols, that is less deserving of protection. This determination would be based on a weighing of costs and benefits that is subjective and arbitrary, may reflect impermissible government motive, and removes a lot of speech from the safeguards of strong First Amendment protection. Then, a court would have to hold that university suppression of a racist symbol survives intermediate scrutiny. This approach may give even less guidance for the future because individuals cannot depend on the fact that a law will generally not survive intermediate scrutiny.

What courts have done in the area of university racist speech is an interesting mix of these approaches. Courts, for good reason, are loath to create an unprotected category of hate speech. When students sue universities for speech codes banning racist speech, lower courts generally strike down these speech codes.[284] However, the Supreme Court has held that universities can be sued for harassment under Title VI or Title IX, but only if racist or sexist speech becomes so severe, pervasive, and both objectively and subjectively offensive that it essentially denies a student her learning opportunity.[285] At that point, and only in severe and pervasive cases, do the speech's vices more closely resemble conduct harms, because the speech directly impacts the material interest of educational benefits.

This approach threads the needle in a case where speech leads to what many might consider an intolerable situation in a unique university context. So long as courts are using a high bar to apply Title VI's harassment standard, this seems like a good use of a hybrid between creating a new category of unprotected speech and simply applying strict scrutiny. Unfortunately, laws against harassment at universities and the workplace, even applying the "severe and pervasive" standard, may restrict too much protected speech, if the bar for when sexist and racist speech rises to the

283. See Smolla, supra note 23, at 511 (describing Justice Breyer's intermediate scrutiny approach in *Alvarez* as "invok[ing] no principled methodology at all, other than to announce that intermediate scrutiny was the proper standard").

284. See, e.g., Dambrot v. Cent. Mich. Univ., 55 F.3d 1177, 1182 (6th Cir. 1995) (invalidating as overbroad "any intentional, unintentional, physical, verbal, or nonverbal behavior that subjects an individual to an intimidating, hostile or offensive educational, employment or living environment" through use of demeaning racial slurs or symbols).

285. See Davis v. Monroe Cty. Bd. of Educ., 526 U.S. 629, 651 (1999) ("[A] plaintiff must establish sexual harassment of students that is so severe, pervasive, and objectively offensive, and that so undermines and detracts from the victims' educational experience, that the victim-students are effectively denied equal access to an institution's resources and opportunities."); see also Saxe v. State Coll. Area Sch. Dist., 240 F.3d 200, 205–10 (3rd Cir. 2001) (detailing analogous standards for Title VI, which prohibits racial discrimination in education; Title VII, which prohibits workplace harassment; and Title IX, which prohibits sexual harassment in education).

level of harassment becomes lower and lower.[286] That, in essence, is the slippery slope problem, and is the reason that harms balancing should occur only in rare cases.

A few points should be stressed. First, courts should avoid creating entirely new unprotected categories of speech, unless exempting speech from protection at that higher level of abstraction would mitigate the possibility of impermissible governmental motives skewing the marketplace of ideas. Sometimes, the only way to avoid viewpoint discrimination, the worst type of First Amendment offense, is to create an unprotected category of speech based on content, or subject matter.[287] This should be avoided if possible, however, because a large amount of speech will then be vulnerable to suppression. Second, courts should conduct Justice Breyer-like balancing only in limited, rare, extreme cases where intermediate scrutiny does not seem arbitrary and likely does not reflect a court's subjective viewpoint preferences, as in the case of commercial speech. To avoid the corrosive effects of free speech consequentialism, many of the harms caused by speech will have to remain unaddressed by law, and dealt with instead using the powerful force of social stigma.[288]

B. *Revenge Porn*

The term "revenge porn" refers to the distribution, without consent, of sexually graphic "images originally obtained *with* consent within the context of a private or confidential relationship."[289] Revenge porn derives its name from the fact that most of the images are posted online in order to

---

286. See Eugene Volokh, Freedom of Speech and Appellate Review in Workplace Harassment Cases, 90 Nw. U. L. Rev. 1009, 1011 (1996) (outlining interaction between First Amendment and harassment claims); see also Eugene Volokh, The Administration Says Universities Must Implement Broad Speech Codes, The Volokh Conspiracy (May 13, 2013, 1:40 PM), http://volokh.com/2013/05/13/the-administration-says-universities-must-implement-broad-speech-codes-2/ [http://perma.cc/F7C8-9R5K] (arguing Department of Justice and Department of Education's recommended speech codes limiting nonsevere and nonpervasive sexual language "extend to speech that is protected by the First Amendment").

287. The difference between content discrimination and viewpoint discrimination is not precise, but viewpoint discrimination is considered the most pernicious subset of content discrimination. If a restriction against lobbying exemplifies content-based discrimination, a restriction against lobbying in favor of animal rights is an example of viewpoint-based discrimination. See Kent Greenawalt, Viewpoints from Olympus, 96 Colum. L. Rev. 697, 698–701 (1996) (surveying Court's treatment of content and viewpoint discrimination).

288. Scholarship in the civil and criminal law area indicates that individuals' behavior may be controlled more by their sense of socially acceptable conduct than actual law. See Edwards, supra note 231, at 50–51 (detailing scholars who have demonstrated "norms of social behavior are often more powerful than the law").

289. Mary Anne Franks, Criminalizing Revenge Porn: A Quick Guide 1, http://alicelaw.org/uploads/asset/asset_file/1979/Criminalizing_Revenge_Porn_Quick_Guide.pdf [http://perma.cc/4TQU-NFUL] (last visited Jan. 23, 2016) (emphasis added).

shame former sexual partners following a breakup.[290] Along with the images, former lovers sometimes post identifying information.[291]

Applying a conduct–harms constraint to free speech consequentialism highlights why and when revenge porn may be regulated. Revenge porn's regulability should stem from the fact that, when the speech was produced, there was an implied sense that it would not be displayed to others. If the targets of revenge pornography knew that images and videos of their intimate activities would be posted publically, the speech never would have been produced in the first place. The regulability of revenge porn thus depends on its connection to a harm that is more akin to conduct—breach of a contract—than from the myriad harms that flow from revenge porn's expressive aspects. Revenge porn can thus be regulated by first presuming the speech is not categorically unprotected, and then considering the speech harms that can be analogized to conduct under a strict scrutiny test.

Revenge porn is designed to cause harm, and it succeeds. The subjects of revenge porn experience not only humiliation, but have also been fired from their jobs, lost friends, and gained the unwanted attention of stalkers.[292] Aggrieved individuals have taken several routes in attempts to have content depicting them removed from the Internet, or to recover damages for their injuries. Targets who took the pictures themselves have sued Internet websites that host revenge porn for copyright infringement or other privacy torts, but federal law immunizes websites from liability based on content posted by users unless the website edits or revises the posted material.[293] Invasion-of-privacy tort claims against

---

290. See Man Charged in Running Revenge Porn Site Convicted, CBSNews (Feb. 2, 2015, 10:04 PM), [hereinafter Man Charged, CBS News] http://www.cbsnews.com/news/man-charged-in-running-revenge-porn-site-convicted/ [http://perma.cc/M34U-BX7D] ("The term revenge porn is used because most of the explicit images have been posted online by former lovers in attempts to shame their former partners after a breakup.").

291. See Eric Schulzke, California Lawmakers Target 'Revenge Porn' but Miss, Critics Say, Deseret News (Sept. 8, 2013), http://www.deseretnews.com/article/865586019/California-lawmakers-target-revenge-porn-but-miss-critics-say.html [http://perma.cc/DLC2-63H9] (explaining revenge porn involves posting pictures "along with identifying information").

292. See Erica Goode, Victims Push Laws to End Online Revenge Posts, N.Y. Times (Sept. 23, 2013), http://www.nytimes.com/2013/09/24/us/victims-push-laws-to-end-online-revenge-posts.html (on file with the *Columbia Law Review*) ("The effects can be devastating. Victims say they have lost jobs, been approached in stores by strangers who recognized their photographs, and watched close friendships and family relationships dissolve. Some have changed their names or altered their appearance.").

293. See Rebecca Tushnet, How Many Wrongs Make a Copyright, 98 Minn. L. Rev. 2346, 2346 (2014) ("At (First Amendment-inflected) common law, it is hard to hold distributors liable for such harm, and Congress has made it impossible online by enacting § 230 of the Communications Decency Act, which provides total immunity to a service provider uninvolved in developing tortious content.").

individuals who post the pictures yield mixed results,[294] and some scholars have even suggested treating revenge porn as a form of Internet sexual harassment.[295]

States have begun to criminalize the publication of nude photos if the person publishing the photos knows or should have known that the subject of the image did not consent to the disclosure.[296] Virginia lawmakers introduced legislation, for example, that would criminalize publishing sexually explicit pictures of someone without permission and with "the intent to cause them substantial emotional distress."[297] A California defendant was convicted of felony charges of identity theft and extortion, for running a revenge porn website where he made aggrieved ex-lovers pay to have their photos removed from his site.[298] His lawyer argued that although his behavior was immoral and offensive, he did not break any laws by allowing others to post sexually explicit photographs.[299]

The regulation of revenge porn presents thorny First Amendment issues, even though the speech is considered both highly injurious and of low value.[300] Some argue that revenge porn can be regulated as obscenity,[301] but, like much pornography, sexually explicit speech that does

---

294. See Daniel J. Solove, The Future of Reputation: Gossip, Rumor, and Privacy on the Internet 119–20 (2007) ("[P]eople suing under the privacy torts frequently lose their case.").

295. See Citron & Franks, supra note 82, at 354 ("Revenge porn is a form of cyber harassment and cyber stalking whose victims are predominantly female.").

296. See Paul J. Larkin, Jr., Revenge Porn, State Law, and Free Speech, 48 Loy. L.A. L. Rev. 57, 94–97 (2014) ("Several states have decided to address [the revenge porn] problem by making revenge porn a crime."); Anne Flaherty, 'Revenge Porn' Victims Press for New Laws, Associated Press (Nov. 15, 2013, 5:34 PM), http://bigstory.ap.org/article/revenge-porn-victims-press-new-laws [http://perma.cc/6X2D-KB7H] ("An increasing number of states, including Maryland, Wisconsin and New York, are considering whether to make it illegal to post any sexually explicit image online without that person's permission.").

297. Lisa De Bode, Outcry Growing Against Revenge Porn, Aljazeera Am. (Jan. 14, 2014, 9:34 PM), http://america.aljazeera.com/articles/2014/1/14/growing-movementagainst revengepornaimstoendonlineharassment.html [http://perma.cc/Z83A-JRFL].

298. See Man Charged, CBS News, supra note 290 ("[V]ictims could go and be charged up to $350 to have the images removed.").

299. Id.

300. See Derek E. Bambauer, Exposed, 98 Minn. L. Rev. 2025, 2031 (2014) (noting greatest value in sexually explicit speech shared between two partners is its ability to "bring[] people, particularly those in intimate relationships, closer together, and allow[] them to express romantic and sexual feelings in new ways"). But see id. (noting value is realized before someone shares material online without partner's consent and undermines speech because "risk of non-consensual distribution or display threatens to undercut the generation of intimate media").

301. See Adrienne N. Kitchen, Note, The Need to Criminalize Revenge Porn: How a Law Protecting Victims Can Avoid Running Afoul of the First Amendment, 90 Chi.-Kent L. Rev. 247, 277 (2015) ("Revenge porn is not protected speech because it is obscene, and therefore falls under a categorical exception to the First Amendment.").

not rise to the level of obscenity is still protected speech.[302] Criminal statutes and torts based on the invasion of privacy and emotional distress caused by revenge porn compromise the freedom to distribute protected speech lawfully obtained. Indeed, the Supreme Court has recognized a right for the media to publish even unlawfully obtained content, so long as the publisher was not involved in the illegal so long as the publisher was not involved in the illegal conduct that produced the content.[303] And in *United States v. Stevens*, the Supreme Court held that individuals cannot be held criminally liable for distributing speech depicting illegal acts, so long as the individuals did not perpetrate the underlying act.[304] Revenge porn, as defined here, is both legally obtained and depicts a legal act.

In the ultimate articulation of free speech consequentialism, Mary Anne Franks argues for criminalization of revenge porn because "some expressions [of free speech] are just considered so socially harmful and don't contribute any benefits to society."[305] Yet this does not separate revenge porn from any number of categories of protected speech that may cause others emotional distress and are considered by some to possess little value; this is nothing more than a call for judges to make wholesale and retail judgments about the value and harms that flow from particular forms of speech. If revenge porn can be regulated, legislators should not target the victim's emotional distress or the invasion of privacy, as these focal points threaten to undermine strong free speech protections exceptional to America's free speech regime.

Privacy cannot be a focal point for regulating revenge porn because America's conception of privacy distinguishes between governmental intrusions on privacy and individual intrusions on privacy.[306] The regulation of individual intrusions on privacy generally must yield to First Amendment concerns.[307] This favoritism of free speech over privacy is

---

302. See Lisa N. Seidman, The First Amendment and Pornography, 64 Geo. Wash. L. Rev. 1052, 1060 (1996) (noting "production of sexually explicit material portraying performers over the age of eighteen" would be protected speech).

303. See Bartnicki v. Vopper, 532 U.S. 514, 535 (2001) (holding First Amendment protects speech disclosing contents of illegally intercepted communication). Although *Bartnicki* involved "matters of public importance," the information in *Bartnicki* was obtained unlawfully. Id. at 533–34. In this Article, revenge porn is obtained lawfully.

304. 559 U.S. 460, 482 (2010) ("We . . . do not decide whether a statute limited to crush videos or other depictions of extreme animal cruelty would be constitutional. We hold only that [the statute] is not so limited but is instead substantially overbroad, and therefore invalid under the First Amendment.").

305. De Bode, supra note 297 (discussing Professor Franks's argument).

306. See James Q. Whitman, The Two Western Cultures of Privacy: Dignity Versus Liberty, 113 Yale L.J. 1151, 1161 (2004) ("America . . . is much more oriented toward values of liberty, and especially liberty against the state. At its conceptual core, the American right to privacy . . . is the right to freedom from intrusions by the state, especially in one's own home.").

307. Id. at 1209 ("Freedom of expression just about always wins in America—both in privacy cases and in cases involving infliction of emotional distress . . . .").

what separates America from other Western democracies.[308] The United States allows individuals wide latitude to publish truthful (yet embarrassing or accusatory) information about each other and preserves spaces of autonomy as against the government far more than against other individuals.[309] The European Union's "right to be forgotten," which allows individuals to sue to have embarrassing articles removed from search engines, could never exist in America.[310] America's approach allows individuals choices as to how they conduct their private interactions with others, separate from the reach of the government.[311]

Further, there is not a principled way to create, at the wholesale level, an unprotected category of speech around revenge porn. The law cannot prohibit the telling of secrets, which are also generally told with an either explicit or implicit understanding that the secret will not be revealed to others.[312] Analogously, the underlying fact that someone took naked pictures should be considered protected speech even if the resulting images are regulable. However, it is likely that couples would still date and engage in intimate behavior even knowing their partners are likely to gossip to friends; it is much less likely that the content of revenge pornography would be produced knowing that it would be publically available. As a result, if the costs of revenge porn so far exceed its benefits that revenge porn fits into the rare case where free speech consequentialism should prevail, states would have to narrowly tailor their laws to target the action of breaching an express or implied promise.[313]

The best way to regulate revenge porn is likely through private (noncriminal) law, as it is the breach of implied contract that should be regulable, not the content of the expression or the emotionally distressing effects of the expression. Using contract principles, courts could award

---

308. See id. ("Freedom of expression has been the most deadly enemy of continental-style privacy in America.").

309. See id. at 1158–59 (noting difference between American and continental approaches to freedom of expression).

310. Editorial, Europe's Expanding 'Right to Be Forgotten,' N.Y. Times (Feb. 4, 2015), http://www.nytimes.com/2015/02/04/opinion/europes-expanding-right-to-be-forgotten.html (on file with the *Columbia Law Review*) (discussing implications of right to be forgotten in United States); see also Paul M. Schwartz, The EU–U.S. Privacy Collision: A Turn to Institutions and Procedures, 126 Harv. L. Rev. 1966, 1995 (2013) (noting right to be forgotten in European Union "raises difficulties regarding the necessary balance of privacy against the freedom of expression").

311. See Whitman, supra note 306, at 1210 ("[N]othing in the doctrine of the 'right of publicity' prevents Americans from alienating the rights in their image, no matter how humiliating their subsequent use may be. . . . In Europe, by contrast, as we have seen, sales of your nude image remain voidable . . . .").

312. Diane L. Zimmerman, Requiem for a Heavyweight: A Farewell to Warren and Brandeis's Privacy Tort, 68 Cornell L. Rev. 291, 293 (1983) (arguing tort of public disclosure of private facts "cannot coexist with constitutional protections for freedom of speech and press").

313. See Larkin, supra note 296, at 104–05 (suggesting narrow tailoring is one way state law can combat revenge porn).

injunctions for specific performance, given that the pictures at issue are unique, the harms at issue do not translate well into market-value damages calculations, and the plaintiff will suffer irreparable injury.[314] Treating revenge porn as a private law issue accords with Solomon and Oman's reasoning that private law concerns should sometimes trump the public law approach of First Amendment jurisprudence better than Solomon and Oman's actual target, IIED[315]—in the case of revenge porn, contract or tort law could allow individuals to set default rules about the speech they produce together in an intimate relationship. This stands in stark contrast to the IIED tort, especially the one at issue in *Snyder v. Phelps*, where the speech is not produced together, and there is no understanding between the two parties that leads to the production or the dissemination of the speech.[316] Further, juries would not decide how "outrageous" the speech was, a content-based judgment about speech, but whether the parties manifested an understanding about the material exchanged between them. Burden-of-proof issues would become important, but evidence of the "revenge" element, or posting after a breakup and with contact information, can be prima facie evidence that one partner knew the pictures produced should remain private. Generally, images that are produced together, or where a partner gives permission to take photographs conditioned on a promise of confidentiality,[317] should be more susceptible to the claim that the revenge porn publisher breached an implied or express promise of confidentiality than "selfies," taken by the subject of the revenge porn and sent to her partner.

Courts should take care not to allow the regulation of revenge porn to create slippery slope problems for other types of confidential information consensually obtained. Courts can focus on the fact that sexually explicit speech, while not categorically unprotected, is sometimes considered of lower value (while being careful not to wholly endorse Justice Breyer's strong form of free speech consequentialism at multiple levels of abstraction outside the sexually explicit speech arena, which is already considered of lesser value).[318] If courts wish to harmonize regulation of revenge porn with a principled free speech jurisprudence, some harms from revenge porn will not have a legal remedy. However, partners can protect their own interests by explicitly asking their partners to agree not

---

314. See Tri-State Generation & Transmission Ass'n v. Shoshone River Power, Inc., 874 F.2d 1346, 1353–54 (1989) (allowing permanent injunction in cases where monetary damages are inadequate to compensate plaintiff).

315. See supra notes 58–66 and accompanying text (discussing Solomon and Oman).

316. See supra note 65 and accompanying text (discussing Snyder v. Phelps, 131 S. Ct. 1207 (2011)).

317. See Citron & Franks, supra note 82, at 345 (detailing true story of "Jane," who "allowed her ex-boyfriend to photograph her naked because, as he assured her, it would be for his eyes only," but whose ex-boyfriend later uploaded her naked photo along with her contact information to popular revenge porn website).

318. See supra notes 165–178 and accompanying text (discussing Justice Breyer's strong form of free speech consequentialism).

to disclose materials in exchange for any photographs.[319] We should also, as a society, address problems in our culture that lead individuals, largely men, to shame and embarrass former intimate partners in this way.[320] Suppressing speech cannot cure the ills that befall both genders in a society still clinging to strong notions of gender and gender roles. It may be true that the Internet is a more hostile place for women, but there may also be spaces, on the Internet or elsewhere, where men feel alienated and more fearful of speaking.

C. *Trigger Warnings*

A movement to include "trigger warnings" on material that may distress students has gained momentum on college campuses.[321] This movement exemplifies why, in line with this Article's proposal, a focus on the emotional harms caused by speech can undermine the value of speech, may ultimately be counterproductive, and often serves as a vehicle for suppression of speech on the basis of viewpoint. First Amendment problems arise if university administrators at public universities require professors to provide trigger warnings for material that might trigger unpleasant emotional responses for students.[322] At public universities, this sort of requirement amounts to viewpoint discrimination and violates professors' academic freedom rights. When this movement impacts policy at private universities,[323] it illuminates important lessons and provides compelling analogies for why courts should account for harms when free speech consequentialism is unavoidable.

Students have begun to lobby college administrations or individual professors for disclaimers that certain books or films discussed in class contain content that depicts sexual violence, racism, sexism, or other

---

319. See Cohen v. Cowles Media Co., 501 U.S. 663, 668–72 (1991) (allowing claim of promissory estoppel against newspaper who agreed to keep source's identity confidential).

320. See Martha C. Nussbaum, Objectification and Internet Misogyny, *in* The Offensive Internet 68, 84–87 (Saul Levmore & Martha C. Nussbaum eds., 2010) (suggesting changes in cultural norms as way of combating revenge porn).

321. See Comm. A on Acad. Freedom & Tenure, On Trigger Warnings, Am. Ass'n of Univ. Professors, [hereinafter Committee A] http://www.aaup.org/report/trigger-warnings [http://perma.cc/CX7F-U89P] (last visited Jan. 23, 2016) (describing recent "demand that teachers provide warnings . . . [for] anything that might trigger difficult emotional responses for students" in colleges).

322. See Erica Goldberg & Kelly Sarabyn, Measuring a "Degree of Deference": Institutional Academic Freedom in a Post-*Grutter* World, 51 Santa Clara L. Rev. 217, 249–50 (2011) (discussing First Amendment academic freedom rights of professors at public universities to choose manner in which they teach classes).

323. The usage of trigger warnings began on feminist websites to caution potential rape victims about graphic depictions of sexual violence in order to avoid triggering a post-traumatic stress disorder response. See Committee A, supra note 321 ("The specific call for 'trigger warnings' began in the blogosphere as a caution about graphic descriptions of rape on feminist sites, and . . . even set off a post-traumatic stress disorder (PTSD) response in some individuals.").

unpleasant topics or events. Advancing the claim that this material may "trigger" students who have been victimized to re-experience a previous trauma, these students have successfully begun to change the culture of learning at universities. Oberlin College, for example, adopted a policy, since tabled due to faculty opposition, which included as possible trigger topics "racism, classism, sexism, heterosexism, cissexism, ableism, and other issues of privilege and oppression."[324] The policy cited Chinua Achebe's novel *Things Fall Apart* as an example of something that might "trigger readers who have experienced racism, colonialism, religious persecution, violence, suicide and more," and suggested that faculty "[r]emove triggering material when it does not contribute directly to the course learning goals."[325] Student leaders at the University of California Santa Barbara, a public university, adopted a resolution that university administrators require trigger warnings on class syllabi, and that they issue advance alerts and permit students to skip class when professors present "content that may trigger the onset of symptoms of Post-Traumatic Stress Disorder."[326]

These actions, if indicative of a trend, may contribute to the deterioration of the intellectual climate on university campuses, where professors are increasingly afraid of promoting debate on topics of significance that may rouse student emotions. Many professors have become so concerned about broaching controversial subjects that they have eliminated essential material from their syllabi. Rape law, for which feminist law professors campaigned to be included in the criminal law curriculum, is increasingly being omitted from professors' syllabi due to fears that professors may say the wrong thing, or that victims of rape will experience grave distress based on insensitive remarks by students.[327] Trigger warnings, and the obsession with the emotional impacts of speech, have already begun to undermine the value of classroom discussion. The American Association of University Professors has noted that "[s]ome discomfort is inevitable in classrooms if the goal is to expose students to new ideas, have them question beliefs they have taken for granted, grapple with ethical problems they have never considered, and, more generally, expand their horizons so as to become informed and responsible democratic citizens."[328]

---

324. Id.

325. Id.

326. Jenny Jarvie, Trigger Happy, New Republic (Mar. 3, 2014), http://www.newrepublic.com/article/116842/trigger-warnings-have-spread-blogs-college-classes-thats-bad [http://perma.cc/N2YE-7H6P].

327. See Jeannie Suk, The Trouble with Teaching Rape Law, New Yorker (Dec. 15, 2014), http://www.newyorker.com/news/news-desk/trouble-teaching-rape-law [http://perma.cc/2LB2-469P] ("Even seasoned teachers of criminal law, at law schools across the country . . . are seriously considering dropping rape law and other topics related to sex and gender violence.").

328. Committee A, supra note 321.

Of course, those who support trigger warnings wish to protect vulnerable students from experiencing not just discomfort, but actual, mental trauma based on diagnosable conditions like posttraumatic stress disorder (PTSD).[329] Proponents of trigger warnings want every student to feel safe in the classroom, and for a multiplicity of identities to be recognized and affirmed.[330] However, the efficacy of trigger warnings has not been systematically studied, and some psychologists actually recommend exposure, not avoidance, as a method of coping with stimuli that are believed to trigger PTSD responses.[331] The list of what may trigger students has grown, and now seems directed at topics that may enrage and upset students, not just because of a traumatic personal experience. Given that there is dispute about the link between reading unpleasant material and experiencing a PTSD response,[332] trigger warnings may actually be a pretextual way of stifling certain viewpoints or prejudicing a debate before it has begun.[333]

Casting a pall over certain materials or viewpoints as exacerbating mental health issues will necessarily stifle certain viewpoints. Indeed, placing trigger warnings on material signals to students that the issues mentioned in the trigger warnings are paramount, and expresses a particular ideology before the students even engage with the material. Many survivors of trauma, who are triggered by particular content in classrooms, believe that trigger warnings are a way to show "understanding and

---

329. See, e.g., Kai Johnson et al., Our Identities Matter in Core Classrooms, Colum. Spectator (Apr. 30, 2015, 1:02 AM), http://columbiaspectator.com/opinion/2015/04/30/our-identities-matter-core-classrooms [http://perma.cc/6WKQ-M3XB] ("The goal . . . was to . . . hold a safe and open dialogue about experiences in the classroom that all too often traumatize and silence students.").

330. See id. ("Students need to feel safe in the classroom, and that requires a learning environment that recognizes the multiplicity of their identities.").

331. See, e.g., Richard J. McNally, Hazards Ahead: The Problem with Trigger Warnings, According to the Research, Pac. Standard (May 20, 2014), http://www.psmag.com/health-and-behavior/hazards-ahead-problem-trigger-warnings-according-research-81946 [http://perma.cc/JP3T-T2M7] ("According to a rigorous analysis by the Institute of Medicine, exposure therapy is the most efficacious treatment for PTSD . . . ."); Florence Waters, Trigger Warnings: More Harm than Good?, Telegraph (Oct. 4, 2014, 2:00 PM), http://www.telegraph.co.uk/culture/books/11106670/Trigger-warnings-more-harm-than-good.html (on file with the *Columbia Law Review*) ("[T]he treatment that we are using to help trauma survivors . . . involves encouraging the patient . . . to make a point of exposing themselves to reminders of trauma so that they can develop a tolerance.").

332. See Jarvie, supra note 326 ("Two people who have endured similarly painful experiences, from rape to war, can read the same material and respond in wholly different ways.").

333. For this reason, student journalists have expressed disapproval of the trigger-warning phenomenon. See, e.g., Ankit Jain & Sarah Manhardt, Fingers Off the Trigger, Gate (Nov. 18, 2014), http://uchicagogate.com/2014/11/18/fingers-off-the-trigger/ [http://perma.cc/XQ8Q-E27F] (declaring "opposition to trigger warnings within all newspapers, because of their subjective nature and potential to be abused to limit free speech").

empathy" to their viewpoints and experiences.[334] Proponents of trigger warnings seem to want their views affirmed, instead of challenged, but topics such as injustice, racism, sexism, oppression, and even rape can be approached from a variety of perspectives instead of being presumed too sacred or offensive to discuss. Students must learn to engage with those who disagree, and should understand that not being affirmed is not the same as being ostracized. Professors should, to some extent, account for the emotional well-being of their students; there is no reason to needlessly distress students for no pedagogical purpose. However, professors should be entitled to employ a diversity of styles to inspire students to perform at their best, and certainly should not make their pedagogical decisions based on fear of healthy, challenging debate. The sanitizing of our classrooms and our discourse presents a major cultural problem.

The answer is not, of course, to have the government intervene to suppress the speech suppressors, but to allow individuals more choice without state intervention. As a private solution to what many consider harmful speech, trigger warnings must be tolerated, although discouraged in light of the lack of indication that they are effective and their propensity to skew debate. Individual professors and even private universities are free to implement whatever policy on trigger warnings they choose, and students experiencing mental health issues should be provided with resources outside the classroom.

But when mandated by administrators at public universities, trigger warnings should be deemed unconstitutional, as a violation of the professors' First Amendment rights of free speech and academic freedom.[335] Although institutions themselves have academic freedom rights to dictate subjects of study and curricula, and these rights often exist in tension with individual professors' academic freedom rights, the sort of suppression of speech inherent in trigger warnings amounts to censorial viewpoint discrimination.[336] Decisions regarding how to run one's classroom to best promote discussion and learning are within the purview of professors' academic freedom rights, and even controversial and inflammatory decisions on how to conduct one's classroom have been left to individual professors.[337]

---

334. See Ponta Abadi, Trigger-Warning Debate Ignores Survivors' Voices, Ms. Magazine: Blog (May 29, 2014), http://msmagazine.com/blog/2014/05/29/the-trigger-warning-debate-ignores-survivors-voices/ (on file with the *Columbia Law Review*) (presenting different testimonial perspectives on trigger warnings).

335. See generally J. Peter Byrne, The Threat to Constitutional Academic Freedom, 31 J.C. & U.L. 79 (2004) (discussing professors' academic freedom rights).

336. See Goldberg & Sarabyn, supra note 322, at 234 ("[A]cademic freedom as a constitutional right should apply [to universities] only when the institutions proffer ideologically neutral goals . . . .").

337. See, e.g., Kathryn M. Stanchi, Dealing with Hate in the Feminist Classroom: Rethinking the Balance, 11 Mich. J. Gender & L. 173, 195–96 & n.55 (2005) (describing sexual harassment cases in which "high value placed on the academic freedom of

The rise of trigger warnings illustrates why the emotional impacts of speech should be downplayed in the consequentialist calculus as compared to more tangible harms. There is no analog to conduct in the harms caused by books and movies discussed in classroom settings. Suppressing or chilling speech with trigger warnings because of its communicative, nonconduct harms necessarily prioritizes the emotions of some over others in an unavoidably subjective way that may serve as a pretext for viewpoint discrimination instead of alleviating tangible harm. The trigger warnings phenomenon also highlights how focusing on one's emotional upset may actually amplify it. If exposure instead of avoidance is actually better for victims of trauma, a culture of open discourse and dialogue will be altered without a corresponding benefit. The art, literature, and films presented in classrooms represent paradigmatic cases of pure speech, where a large part of their value stems from the ability to provoke new feelings and emotions, and create cognitive dissonance that leads to change. Administrators forcing trigger warnings on professors, in an attempt to nullify the effects of triggering speech, will impose a sterile environment, where feelings are elevated over discourse.

D. *Balancing in Tort Versus Criminal Law*

Finally, as mentioned in Part II, scholars who wish to better incorporate harms into free speech jurisprudence often propose to do so using tort law, which allows for greater flexibility in balancing the harms and benefits of speech.[338] This flexibility, however, mirrors the threats to principled free speech jurisprudence that arise from a greater incorporation of free speech consequentialism. For several reasons, courts should be careful not to elevate tort law as a means to allow restrictions on speech not permitted via criminal law for several reasons.

First, there is insufficient support for Sacks's and others' claims that tort law will not chill speech that does not cause actual harm.[339] Negligence law, precisely because of the flexible outcomes of its tests for foreseeability and proximate causation, is necessarily less predictable. Constitutionalized negligence, as Sacks proposes, would allow individual juries to examine the costs and benefits of speech at the retail level, or, at the least, would allow courts to decide as a matter of law if public policy favors penalizing the producers of what Sacks calls "unreasonably dangerous speech," like violent video games, such that they internalize the costs of the violence correlated with their sales.[340] In assessing tort liability for unreasonably dangerous speech, Sacks proposes that courts evaluate the

professors has led courts to permit [inappropriate] professorial behavior in the classroom").

338. See supra notes 46–63 and accompanying text (surveying theory of tort law application in First Amendment jurisprudence).

339. See supra notes 53–57 and accompanying text (discussing Sacks's theory).

340. Sacks, supra note 15, at 1084.

nature of the speech, the vulnerability of the plaintiff, and the state's interest in punishing the speech to determine the evidentiary burden placed on plaintiffs.[341] This could vary by jurisdiction and would substantiate the fears of unprincipled, subjective jurisprudence that free speech consequentialism inspires. The incitement standard, combined with criminalizing actual violence, seems a wiser approach to managing the harms of advocacy.

Moreover, Oman and Solomon may be generally correct that private law should not be considered simply a species of public regulation, and that tort jurisprudence should generally focus more on individual justice than be viewed as public regulation.[342] However, in the First Amendment context, this approach to speech-torts would be unduly corrosive to free speech rights. It would neglect the fact that state-imposed penalties for speech-torts can exert a powerful chilling effect on speech. Unlike regulating revenge porn civilly as an implied breach of contract[343]—where juries can be charged to decide what sort of arrangement intimate partners had with respect to the speech produced—cases of IIED or negligence for unreasonably dangerous speech empower citizens, with the legal power of the state, to target messages they dislike. Even if, as Oman and Solomon note, "[t]he outrageousness requirement [of an IIED tort] is to make sure that the speech is sufficiently egregious that it is not simply something that the majority doesn't like," and even if the plaintiff is not really acting on behalf of the state, like some sort of private attorney general,[344] the government, at a higher level of abstraction and generality, has empowered citizens to target messages that are offensive at a low level of abstraction. The government may not be systematically rooting out viewpoints it dislikes, but juries can use their charge to suppress unpopular views, which may be those that need the most protection.

Oman and Solomon's argument that the state interest in IIED torts should be recast as primarily giving plaintiffs a mechanism for redressing harm rather than suppressing speech may be accurate.[345] But their solution—to tinker with speech torts to allow, on a case-by-case basis, "a more nuanced assessment of First Amendment interests in light of the

---

341. See id. at 1101, 1103–04 (suggesting Supreme Court's "evidentiary tailoring method of constitutionalizing speech-torts . . . is instructive on how courts could constitutionalize the tort of negligence relative to unreasonably dangerous speech").

342. See id. at 1114–24 (discussing competing views of private law as form of public regulation versus instrument of individual justice).

343. See supra notes 313–317 and accompanying text (discussing revenge porn through implied breach of contract lens).

344. Oman & Solomon, supra note 58, at 1140–41.

345. See id. at 1154–55 ("Although the state may be interested in suppressing certain kinds of wrongs that give rise to torts . . . the primary purpose of tort law itself" is to "provid[e] citizens with recourse against those that have harmed them.").

state interest"[346]—leads to all the problems inherent in Justice Breyer's approach in *United States v. Alvarez*.[347] Oman and Solomon propose that courts perhaps award only actual damages in IIED cases involving speech, or create standards like those in defamation, which focus on the identity of the plaintiff and the purpose of the litigation to determine whether there is a risk of an impermissible speech-suppressing motive.[348] This move combines a purposivist concern for viewpoint suppression with a stronger incorporation of free speech consequentialism, but then loses the benefits of either. Ex ante, it will be difficult for any particular defendant to know whether her speech will be found "outrageous" by a jury that may simply be hostile to her views.

A penalty-sensitive approach to the First Amendment, using tort law instead of criminal law, does have its virtues, such as transparency.[349] If judges, sub silentio, do account for the severity of a sanction when determining a law's constitutionality, it is perhaps better to do so openly.[350] Because of tort law's flexibility, however, there is also a risk of jurors rendering rulings, sub silentio, based on the viewpoint of the speech, not on the harm it actually causes. The application of criminal law may allow for more transparency in cases where criminal law is less vague, with fewer nebulous standards incorporating policy judgments, than in tort law. More importantly, the selection of tort law over criminal law would actualize a strong incorporation of free speech consequentialism that may ultimately prove corrosive to our strong, rules-based First Amendment regime. Courts should thus continue to treat tort law and criminal law more or less indistinguishably when it comes to assessing First Amendment claims unless the speech has conduct-like properties that warrant a distinction between civil and criminal law.

CONCLUSION

Free speech consequentialism, even among those who oppose it, inevitably influences the constitutional protection afforded to particular speech. However, not all free speech consequentialism is created equal. This Article proposes ways to mitigate the subjectivity concerns inherent in balancing harms versus benefits while maintaining strong free speech protections. Courts should consider only those speech harms that can be

346. Id. at 1162.

347. Indeed, Oman and Solomon praise Justice Breyer's concurrence in *Snyder v. Phelps* for its "case-by-case approach to balancing First Amendment and tort interests." Id. at 1135.

348. See id. at 1162, 1166–67 (discussing how remedies to IIED need not be binary, all-or-nothing propositions).

349. See Coenen, supra note 119, at 1042–44 (extolling virtues of "openly acknowledging the First Amendment significance of a penalty's severity").

350. See id. (arguing judges and juries already undertake "penalty-sensitive free speech review . . . behind closed doors," and increasing transparency of this reality would promote understanding of "actual contours of the law").

analogized to harms generally caused by conduct in their consequentialist calculus. This approach constrains free speech consequentialism at all levels of abstraction, or stages of First Amendment jurisprudence, while allowing emerging harms caused by speech to be regulated.