Catholic University Law Review

Volume 71
Issue 3 *Summer 2022*

Article 7

6-20-2022

The Kids are all Right: The Law of Free Expression and New Information Technologies

Mark Tushnet
Harvard University, mtushnet@law.harvard.edu

THE KIDS ARE ALL RIGHT: THE LAW OF FREE EXPRESSION AND NEW INFORMATION TECHNOLOGIES

Mark Tushnet[+]

Recently the literature on free expression has turned to the question, should the law of free expression be adjusted because of the availability of new information technologies (hereafter NIT), and if so, how? The only thing about NIT that distinguishes them from traditional media is that disseminating expression via NIT is much less expensive than doing so via traditional media. The tenor of recent scholarship on NIT and free expression is that the invention of NIT does support some modification of free expression law. This Essay argues that that conclusion might be correct, but that many of the arguments offered in support move much too quickly. The Essay tries to slow them down and in so doing to suggest that the arguments require complex and contestable judgments about exactly how an expansion of expression might elicit new rules regulating it. My goal is to identify the lines of analysis that need to be pursued before we conclude that the existing law of free expression should be modified in response to NIT. In that somewhat limited sense the Essay is contrarian.

The Essay's overall theme is that the First Amendment rules in place probably already accommodate the concerns that motivate arguments for adjusting the law of free expression. The First Amendment, that is, is not obsolete.

[+] William Nelson Cromwell Professor of Law emeritus, Harvard Law School. This Essay develops ideas I sketched in a keynote speech to a conference on free expression sponsored by the International Association of Constitutional Law in July 2020. I thank Rebecca Tushnet and participants in the conference for their comments on prior versions (and note that most of the comments were critical of the position I take here).

INTRODUCTION 472
I. IS THE FIRST AMENDMENT OBSOLETE?: AN ANALYSIS AND CRITIQUE 474
II. THE SPEECH-CAUSES-HARM PARADIGM AND NIT 478
III. NIT AND THE FOUNDATIONS OF FREE EXPRESSION LAW 481
IV. HOW MORE SPEECH CAUSES MORE HARM, AND WHAT MIGHT BE DONE ABOUT IT 482
A. The Implications of a Proportionality Rule in Place 482
B. False Statements That Injure Reputation 485
C. Expression that Induces Law Violation 487
D. Sexually Explicit Material 489
E. False Statements That Cause Systemic but Not Material Harm: Herein of Fake News 492
F. Threats: Herein of Material Harm, Distress and Anxiety, and Cyberstalking 495
V. REGULATING PLATFORMS 499
A. Platforms as Governments 499
B. Platforms and General Laws 500
C. Platform Liability for Utterances Not Protected by Free Expression 501
1. Accomplice Liability 501
2. Incentive Effects 503
CONCLUSION 505

INTRODUCTION

Recent literature on free expression has turned to the question: Should the law of free expression be adjusted because of the availability of new information technologies (hereafter "NIT"), and if so, how?[1] As far as I can tell, the only thing that distinguishes NIT from traditional media is that disseminating expression via NIT is much less expensive than doing so via traditional media.[2]

1. For present purposes the term "new information technologies" refers to internet-based methods of disseminating information, including Facebook, Twitter, TikTok, and WhatsApp. For two recent contributions, *see* Toni M. Massaro & Helen Norton, *Free Speech and Democracy: A Primer for 21st Century Reformers*, 54 UC DAVIS L. REV. 1631 (2021) (addressing the values of the First Amendment—that more speech is always better than less—in a modern environment flooded with potentially harmful speech); Tim Wu, *Is the First Amendment Obsolete?*, *in* THE PERILOUS PUBLIC SQUARE: STRUCTURAL THREATS TO FREE EXPRESSION TODAY 15 (David E. Pozen ed., 2020) (commenting on how NIT are likely to force free speech doctrines to grapple with the need to promote a healthier market place of ideas). *See also* Massaro & Norton, *supra* note 1, at 1638, n.13 (citing earlier versions of the argument).

2. *See, e.g.*, Massaro & Norton, *supra* note 1, at 1639 (referring to "the cheapness of speech"); Wu, *supra* note 1, at 16 (referring to the "low costs of speaking"). Other authors have summarized the "unique" characteristics of NIT as velocity, virality, anonymity, homophily, and monopoly. Nathaniel Persily, *The Internet's Challenge to Democracy: Framing the Problem and Assessing Reforms*, *in* ONLINE PLATFORMS AND LEGAL TECHNOLOGIES 2019: LEGAL AND REGULATORY RESPONSES TO TECHNOLOGY CHALLENGES 585, 5–6 (2019). That these are "unique" or analytically separate from cheapness seems questionable to me. As to velocity and

More precisely, the cost of disseminating a message to an intended audience is much lower.[3] The lower cost leads to more expression, which is said to make greater demands upon the scarce resources of listeners' attention.[4]

The tenor of recent scholarship on NIT and expression is that the invention of NIT does support some modification of current free expression doctrine so as to continue to advance the values protected and promoted by that doctrine. This Essay argues that the conclusion might be correct but that many of the arguments offered in support of it move much too quickly. The Essay tries to slow them down—or, to use a different metaphor, to put more flesh on their bones—and in so doing to suggest that the arguments require complex and contestable judgments about exactly how an expansion of expression might elicit new rules regulating it. The goal is to identify the lines of analysis that the literature should pursue before concluding that the existing law of free expression should be modified in response to NIT, though without tracking every twist and turn in each line. In that somewhat limited sense, the Essay is contrarian.

The Essay proceeds as follows. Section I examines and critiques Professor Tim Wu's prominent version of the argument that the development of NIT should lead us to rethink the law of free expression. Section II presents the paradigm underlying free expression law: that speech causes harm. That paradigm means that the law of free expression attempts to jointly optimize speech and harm by keeping the harm caused by speech to an acceptable level without eliminating that harm entirely. An increase in the amount of expression might lead us to revise free expression law because—simplifying the argument substantially—the more speech, the more harm, which might lead us to seek a new set of rules that once again balance speech and harm.

Section III identifies two kinds of supporting arguments. First, it is said, NIT should lead us to alter substantive First Amendment law and impose liability for utterances that traditionally have received constitutional protection. The general

---

virality, consider the well-known saying, "[a] lie can travel halfway around the world while the truth is putting on its shoes," which regularly comes up in discussions of libel law. *A Lie Can Travel Halfway Around the World While the Truth is Putting on Its Shoes*, QUOTE INVESTIGATOR (July 13, 2013), https://quoteinvestigator.com/2014/07/13/truth/ (discussing the saying's origins). NIT scale up this familiar problem, and for that reason this Essay's focus is on scale or amount of expression. I discuss homophily in Section III and monopoly in Section V(A). NIT are contrasted with categories of "traditional media," including face-to-face interactions, pamphlets, newspapers, movies, radio, and television, and combinations of these media such as videotapes and CDs.

3. Sometimes the effect is to disseminate the message to a wider audience than the speaker cares to reach, but discussions of NIT tend to assume either that the speaker does not care that the message will reach a group larger than its intended audience or that the speaker actually welcomes that effect.

4. *See, e.g.*, Massaro & Norton, *supra* note 1, at 1639 (asserting that "the [sheer] volume of speech . . . now leaves listeners' attention . . . increasingly scarce"); Wu, *supra* note 1, at 23 (referring to "scarcity of attention"). I address the "scarce attention" claim specifically in Section I but thereafter refer only to the increased amount of speech via NIT; such references should be read to incorporate scarce attention.

balance we have struck among the values promoted by free expression has to be reconsidered in light of NIT.

The second argument identified in Section III is that NIT is said to affect the mechanisms by which specific categories of speech cause specific harms. Section IV examines this argument by addressing specific doctrinal areas in a more granular fashion. It looks at the mechanisms by which more speech might render no longer socially optimal rules—what will be called the "rules in place"—that jointly optimize speech and harm in a world with the traditional media only. This Section also looks at false statements that injure reputation (libel); expression that induces unlawful action (the subject of the traditional law of sedition); sexually explicit expression (obscenity and pornography); false statements that inflict no material harm (fake news); and threats (cyberstalking).

Section V turns to arguments about the platforms which make up NIT: Twitter, Facebook, and other social media platforms. These arguments come in several forms. First, the platforms should be subject to the same limitations on speech regulation that apply to the government. Second, the platforms can be regulated through the application of antitrust or fiduciary law without violating the First Amendment. This Section discusses existing First Amendment doctrine about the application of "general" laws to the media and examines some issues that might arise in connection with tinkering with antitrust or fiduciary law as the vehicle for platform regulation. Third, the platforms should be held liable for the utterances they disseminate, holding constant the substantive rules of libel, threats, and other torts involving intentional falsehoods. The overall theme in Section V is that the First Amendment rules in place probably already accommodate the concerns that motivate arguments for the three types of platform regulation.

I. IS THE FIRST AMENDMENT OBSOLETE?: AN ANALYSIS AND CRITIQUE

Professor Tim Wu opens his essay, *Is the First Amendment Obsolete?*, with the observation that "[t]he most important change in the expressive environment can be boiled down to one idea: it is no longer speech itself that is scarce, but the attention of listeners. Emerging threats to public discourse take advantage of this change."[5] Contemporary censorship "targets listeners or [] undermines speakers indirectly" "[i]nstead of targeting speakers directly."[6] It does so through both "new punishments, like unleashing 'troll armies' to abuse the press and other critics, and" by "flooding tactics . . . that distort or drown out disfavored speech."[7] Wu identifies "three developments" that combine to change the information environment: (1) "the massive decrease since the 1990s in the costs of being an online speaker"; (2) "the rise of 'attention industry' . . .

5. Wu, *supra* note 1, at 15.

6. *Id.*

7. *Id.* at 15, 24–25 (describing both "online harassment and attacks" and "distorting and flooding" as "cheaper . . . alternatives" to "direct censorship").

actors whose business model is the resale of human attention"; and (3) "the rise of the 'filter bubble[,]' . . . [meaning] the tendency of attention merchants . . . [to] offer[] audiences a highly tailored, filtered package of information designed to match their preexisting interests."[8]

Wu offers a number of suggested new laws, the constitutionality of which might require "doctrinal evolution"—referred to here as adjustment of the rules in place.[9] These include laws "designed to better combat the specific problem of troll-army style attacks on journalists or other public figures; [and] [n]ew . . . restrictions on the ability of major media and internet speech platforms to knowingly accept money from foreign governments attempting to influence American elections."[10]

Notice first that the proposed law regarding foreign contributions would apply to "major media" as well as internet speech platforms. Wu's inclusion of the major media suggests that the problem does not arise from NIT, although NIT may have brought the problem to our attention. Further, it seems quite clear that the proposed statute is constitutional under the rules in place, so no doctrinal evolution is required.[11]

Professor Wu also suggests that we might have to expand the notion of accomplice liability to deal with the troll armies that public officials can unleash. He offers examples, such as: "the president . . . name[s] individual members of the press and suggest[s] they should be punished, yielding a foreseeable attack; [or] . . . the president . . . order[s] private individuals . . . to attack or punish critics of the government."[12]

The latter example appears to involve a troll army that is indistinguishable from a real army: its members respond to "orders" from the president. It is

---

8. *Id.* at 21. It is worth noting that new technologies have regularly elicited concerns that they change everything: yellow journalism compared to established professionalized or partisan newspapers; William Randolph Hearst as a new kind of media monopolist influencing public policy; or Father Charles Coughlin's use of radio to promote demagogic policies, "I'm mad as hell" from the movie "Network" regarding a televised version of talk radio. Each time it has turned out that existing doctrine was reasonably good at dealing with the problems critics associated with these phenomena.

9. *Id.* at 35.

10. *Id.* at 38.

11. The Supreme Court has affirmed a lower court decision upholding a statutory ban on contributions by non-citizens to U.S. political campaigns. Bluman v. FEC, 800 F. Supp. 2d 281, 288 (D.D.C. 2011), *aff'd*, 565 U.S. 1104 (2012). It has upheld the constitutionality of a statute requiring domestic distributors of certain foreign-produced materials to label them as "political propaganda." Meese v. Keene, 481 U.S. 465, 467, 484–85 (1987). And it has held that the First Amendment does not cover the expressive activities of foreign nationals. Agency for Int'l Dev. v. All. for Open Soc'y Int'l, 140 S. Ct. 2082, 2086 (2020) (stating in a case involving free expression, "it is long settled as a matter of American constitutional law that foreign citizens outside U. S. territory do not possess rights under the U. S. Constitution"). These cases make it clear that the unadjusted rules in place would readily support the constitutionality of a ban on foreign interference in U.S. elections via NIT.

12. Wu, *supra* note 1, at 36–37.

doubtful that anyone would think that a president who ordered soldiers to beat up critics because they were critics had not violated the First Amendment, though there might be problems in devising an appropriate remedy for the violation. Again, no adjustment in the rules in place seems required.

The same may be true of the president who suggests that named individuals should be punished, which foreseeably occurs. First, note that the problem is not new: it is presented by Henry II's purported statement in 1170, referring to Thomas Becket, "Will no one rid me of this meddlesome priest?"[13] Second, consider a milder variant: a hypothetical government-distributed list of works that *no patriotic citizen should purchase.* This government action is recognizably an act of censorship subject to ordinary First Amendment scrutiny.[14]

Finally, it is not clear that accomplice liability for foreseeable unlawful actions requires the person subject to liability to: (1) foresee that the unlawful action will occur—a subjective standard—or; (2) allow punishment if the person reasonably should have foreseen that it would occur—an objective standard—, either in general or in the First Amendment context.[15]

The example of a troll paramilitary raises another concern. Suppose that the president cannot order trolls to turn out. Rather, the soldiers in the troll paramilitary are ordinary citizens who happen to agree with the president's disparaging remarks about journalists.[16] As Geoffrey Stone observes in his comment on Professor Wu's essay, "[i]ndividuals who take positions that offend others . . . have always been vulnerable to condemnation by others, and such personal condemnation by neighbors, friends, employers, and coworkers might be far more daunting that mass condemnation by strangers, especially once people get used to such behavior. And, yet, we lived with it," at least where the condemnation does not come in the form of a true threat.[17] Again, the problem is not new. The only question is whether this sort of condemnation occurs should

---

13. For a discussion of the phrase's genesis, see Emily Guerry, *"Will No One Rid Me of This Meddlesome Priest?": The Truth Behind Henry II's Notorious Lament*, HISTORY EXTRA, Apr. 20, 2021, available at https://www.historyextra.com/period/medieval/will-no-one-rid-me-of-this-meddlesome-priest-truth-henry-ii-quote/.

14. We might think that the rules in place do not strike the right balance because we should put up with more harm than the rules in place allow. On some accounts that invoke institutional limitations on legislative and judicial competence, individuals may think that the socially acceptable level of harm is higher than that generated by the rules in place. If so, the increased injury wrought by NIT might lead us to protect *more* expression to maintain the same (inadequate) level of damage.

15. For further discussion regarding accomplice liability, *see infra* Section V(C).

16. At one point Professor Wu describes the troll armies as "the government's allies." Wu, *supra* note 1, at 27 (emphasis omitted). This is a weaker formulation than the earlier one in which the troll armies took orders from the president, but it raises the same issues about the choice between a subjective and an objective standard of foreseeability for purposes of accomplice liability.

17. Geoffrey Stone, *Reflections on Whether the First Amendment is Obsolete*, *in* THE PERILOUS PUBLIC SQUARE, STRUCTURAL THREATS TO FREE EXPRESSION TODAY 44, 46–47 (David E. Pozen, ed., 2020).

lead us to adjust the rules in place. This question is discussed in Section V(E) below.

A more pervasive concern is that Professor Wu has too narrow a conception of what regulations fall within the First Amendment's domain. His paradigm example is an action that "directly punish[es] disfavored speakers" or "coercive government action"—in essence, locking critics up or fining them.[18] This theory is why he treats the acts of troll armies as a *new* form of punishment and grounds his suggestion that "the First Amendment must broaden its own reach to encompass new techniques of speech control."[19] This is an inadequate description of the First Amendment's domain. Government actions that increase the cost of operating an expressive business are subject to First Amendment scrutiny. Taxes imposed on the press are an obvious example, and not because failure to pay the tax might be punished *coercively*.[20] Similarly, with the list of disapproved publications. Once we see that the punishments Professor Wu describes as new are actually well-established objects of First Amendment concern, we might reevaluate his claim that the "new" punishments require an adjustment of the rules in place. These new punishments might or might not require an adjustment, depending on what the relevant rule is, but that calls for a more discrete analysis than Professor Wu offers.

Finally, what of the flooding of listeners' attention? Here Professor Wu's analysis is incomplete. Start with traditional media. Users have strategies for consuming them. For newspapers, one person's strategy might be to skip the business section, skim the international news, and read stories about baseball but skip stories about hockey. For radio, a strategy might be to listen to one channel until you hear the weather report, then switch to an all-news channel for the headlines, then turn to a talk radio channel. Another example might be a newspaper reader who chooses to read the *Washington Post* rather than the *Washington Examiner*. User strategies reinforce and perhaps intensify the reader's preexisting views. Professor Wu does not discuss why such strategies are unavailable to users of NIT.[21]

---

18. Wu, *supra* note 1, at 31, 34.

19. *Id.* at 32.

20. *See, e.g.*, Minneapolis Star Tribune, Inc. v. Comm'r, 460 U.S. 575, 591–92 (1983) (holding unconstitutional a state tax on the paper and ink used by large-circulation newspapers).

21. Some NIT business models try to keep users on their sites for as long as possible, and certainly have the potential to do so longer than traditional media do: A reader can actually finish a newspaper, but you cannot *finish* YouTube. You can, however, get tired of clicking on seemingly enticing new offerings—and can devise general strategies for avoiding chasing every new rabbit that the site runs across your screen.

I can report only my own experience as a relatively light user of NIT who accepts default privacy settings. I find it a little creepy that shortly after I do a Google search for some product I start getting ads for products in the same or a related category, but I do not have any difficulty skipping them. When bored, I sometimes click on obvious click-bait postings, but can ignore them in the ordinary course. And I undoubtedly have other strategies that operate below the level of consciousness.

As mentioned, Professor Wu refers to "attention merchants" who present information "tailored" to the listeners' "preexisting interests."[22] Why this is a problem is unclear. If the tailoring is perfect, the material delivered cannot distort the listeners' views; it can only reinforce and perhaps even intensify them. These effects are not distinctive to material delivered via NIT.[23] Social psychologists call this *confirmation bias*, and they have studied it for generations, starting well before the arrival of NIT. If the tailoring is imperfect, the material might change the contents of the basket of information users have, but the information merchants' business model will not work well if they try to replace a large portion of the basket with something completely different.

I share Professor Wu's sense that "filter bubbles" are more common today than in the past and that they are more difficult to burst.[24] I am not sure that we are correct in our assessment, though, because we do not have good historical studies of, for example, whether the readers of the intensely partisan newspapers of the nineteenth century lived in filter bubbles too. Even if we are correct, it is not clear that epistemic closure results from NIT rather than from political developments. For example, political polarization drives epistemic closure, which NIT then reinforces through filter bubbles—rather than filter bubbles driving political polarization.[25]

In Professor Wu's work, and in the literature of which it is a part of its own epistemic bubble, assumptions about both existing law and the distinctively harmful effects of expression via NIT go unquestioned. The remainder of this Essay examines many of those assumptions, sometimes finding them warranted and other times finding them more problematic.

II. THE SPEECH-CAUSES-HARM PARADIGM AND NIT

Why should an expansion of expression lead us to reconsider the shape of free expression doctrine? The proposition that lies at the heart of free expression law is that speech causes harm.[26] Many of the arguments for adjusting free

22. Wu, *supra* note 1, at 21.

23. For example, we know that political radicalization has occurred through the dissemination of videotapes, CDs, and through face-to-face discussions ("brain-washing") when groups recruit new members by isolating them from outside contact.

24. Wu, *supra* note 1, at 21–23.

25. Without claiming that social scientists have pinned down how and particularly when the polarization we are experiencing began, some prominent works date it from the 1980s and early 1990s, before NIT became pervasive. *See, e.g.*, THOMAS E. MANN & NORMAN ORNSTEIN, IT'S EVEN WORSE THAN IT LOOKS: HOW THE AMERICAN CONSTITUTIONAL SYSTEM COLLIDED WITH THE NEW POLITICS OF EXTREMISM 31, 44–45, 49 (2012); JULIAN E. ZELIZER, BURNING DOWN THE HOUSE: NEWT GINGRICH, THE FALL OF A SPEAKER, AND THE RISE OF THE NEW REPUBLICAN PARTY 285, 300 (2020).

26. Frederick Schauer has insisted on this proposition throughout his scholarship on the law of free expression. For a recent version, *see* Frederick Schauer, *Rights, Constitutions, and the Perils of Panglossianism*, 38 OX. J. LEGAL STUD. 635, 647–48 (2018). *See also* Mark Tushnet, ADVANCED INTRODUCTION TO FREEDOM EXPRESSION § 1.3 (2018).

expression law in light of NIT take for granted that because more speech causes more harm, an increase in expression could readily justify changing the pre-existing law.[27] In its barebones form, that assumption is mistaken. Adding flesh on the bones, however, turns out to be a complex enterprise.

Free expression doctrine seeks to achieve the level of regulation that jointly optimizes or, in less formal terms, strikes the right balance between expression and harm.[28] The doctrine asserts that the harm caused by constitutionally protected expression is acceptable within the relevant society and that rules protecting expression that causes less harm would be unacceptable, as would the imposition of liability on expression that causes more harm.[29] As a result, the rules in place today, having been developed with traditional media in mind, protect expression that causes an acceptable level or distribution of harm.[30] If speech causes harm, then the more speech there is the more harm there is and the distribution of that added harm might be troubling. Applying the current rules in a world where there is more expression, and in turn, more harm or a problematically different distribution of harm can produce a greater-than-acceptable level or distribution of harm. The argument for changing the rules in place for an NIT world is that some alternative rules can be devised that would reproduce the acceptable level and distribution of harm.[31] NIT means more expression; more expression means more harm. The rules in place are predicated on jointly optimizing expression and harm and may not continue to do so in the NIT world.[32]

---

27. As Geoffrey Stone observes in his comment on Wu, "there was a time in our history when technological advances in the communications market caused similar angst," referring to the development of radio broadcasting. Stone, *supra* note 17, at 46. Because the rules in place at the time imposed few limits on government regulation of expression, they did not need to be adjusted to deal with the then-new technology. Over time the rules applicable to broadcasting have come into line with those applicable to the media that pre-existed radio. Perhaps we should take the overall story about regulating radio as a new technology that gradually became a familiar one as a cautionary tale.

28. Another way of stating this, is that the rules in place identify the point on the Pareto frontier for speech and harm that the society finds acceptable. (In the present context the Pareto frontier is the set of combinations of speech and harm with the characteristic that increasing one element, such as speech, necessarily decreases the other.)

29. Social acceptability includes consideration of the way in which harm is distributed among the population.

30. "Acceptability" is a social judgment, and one nation's legal and social culture at one point in time might find acceptable a doctrinal structure that another nation's culture would find unacceptable, or that the same nation's culture would find unacceptable at another point in time.

31. Again, an alternative rule, or rules, would return us to the point on the Pareto frontier that we, as a society, had previously concluded was where we wanted to be. In this Essay, my shorthand formulations generally refer only to level of harm, but the full formulation would need to take into consideration the distribution of harm.

32. Embedded in this argument is the assumption that the existence of NIT has not changed what the relevant society regards as an acceptable level of harm. Specifically, the existence of NIT is not in itself a reason for increasing or decreasing the socially tolerable amount of harm caused by expression. In the end, we might be unable to devise administrable new rules that keep the harm

So far, the assumption is that more expression means more harm or harm distributed differently. However, it also means more of whatever it is that makes expression socially valuable. The increased social value of increased expression might offset the increased harm or the changed distribution of harm, in which case the rule in place might still optimize expression and harm. Suppose that a lawmaker reasonably concludes that increases in the amount of expression have diminishing marginal value, with each new expression increasing social value but less than the previous increase, *and* that each new expression increases harm by the same amount. For example, there might be some social value in disseminating lies, but the more a lie is repeated, the less each repetition adds to society. Yet, each additional lie might undermine public confidence in the accuracy of *all* statements the public encounters as much as the first or subsequent lie did. The rule in place might then not optimize expression and harm. This possibility lies at the heart of what follows.

The remainder of this Essay examines two ways that the rules in place might be altered.[33] First, NITs might lead us to change doctrine by changing our understanding of the proper balance among the values underlying the rules in place. The second way is by identifying how more speech causes more harm with respect to specific categories of expression. The next Sections of this Essay deal with those methods in succession. However, this Essay focuses on how the structures of free expression doctrine are created, not on what the new rules for NIT should be—it is methodological rather than prescriptive.[34]

Embedded within the argument that NIT should affect the law of free expression because NIT increases the harm caused by speech is a question about the structure of the *new* law of free expression. Should the traditional media continue to be regulated by the rules in place and new rules be developed that are applicable only to NIT? Alternatively, should the new rules be applied to both traditional media and NIT? We can imagine a doctrinal structure where the law responds to the incremental harm caused by each medium: one rule for expression disseminated face-to-face, another for expression over radio, and yet

---

caused by free expression to the level the rules in place do for expression via traditional media. In that case, we would have to accept a new, higher level of harm. I suspect that this constraint might not exist in connection with all expression disseminated via NIT but might exist in connection with some discrete categories of expression. I take a category-by-category approach to the larger question in what follows but do not examine in detail the question of whether we can devise an administrable new rule for NIT with respect to each category.

33. The focus of this Essay is on the structure of the substantive law of free expression. I do not address even broader questions about the political economy of expression in a world with NIT, because the connection between political economy and specific doctrinal formulations is essentially indeterminate (both for the world in which there were only the traditional media and for today's world). *See* JULIE E. COHEN, BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTION OF INFORMATIONAL CAPITALISM 10 (2019) (examining such questions).

34. However, I do sometimes suggest how a rule-structure that features a general preference for proportionality arguments might respond to NIT differently than the way in which more categorical rule-structure, such as that used in the United States, might respond.

a third for expression through NIT. The logic is clear. Each medium adds further harms on the margin, compared to a baseline set perhaps by face-to-face communication, and joint optimization is best achieved by addressing each medium's incremental contribution to harm.

In Section V of this essay, the inquiry is framed in those terms. For example, it asks whether we would penalize a large-circulation newspaper for publishing a story that we would treat as constitutionally protected when published by a small-circulation one. As a matter of the positive law of free expression in the United States and elsewhere, the rules are the same for all traditional media. The reason is likely administrability: defining rules calibrated according to a medium's incremental effect on harm is probably beyond the capacity of courts devising constitutional doctrine.[35]

III. NIT AND THE FOUNDATIONS OF FREE EXPRESSION LAW

A catalog of values defines the contours of freedom of expression: personal autonomy, discovery of truth, promoting the functioning of democracy, suspicion of government's ability to determine the socially acceptable level and distribution of harm, social order, privacy, and equality.[36] These values help us identify the precise contours of the constitutional rules we think will jointly optimize expression and harm.

NIT might make some of these values more vivid. The increased amount of expression might clarify for us the fact that false statements that affect reputation can thereby trench upon privacy interests.[37] The proliferation of sexually explicit material over the Internet might make it clearer that such expression has adverse effects upon women's social standing. We might reconsider the proposition that the rules in place jointly optimize expression and harm.[38]

---

35. The issue of administrability surfaces in discussions of whether "citizen journalists" or bloggers fall within mostly statutory protections afforded to "real" journalists. *See* IAN CRAM, CITIZEN JOURNALISTS: NEWER MEDIA, REPUBLICAN MOMENTS, AND THE CONSTITUTION 119–43 (2015), for a discussion of how several jurisdictions address this question.

36. How equality is built into the doctrine implicates aspects of the state action doctrine that I do not address in this essay. A shorthand version of the state action issue is this: Expression might contribute to inequality by inducing private parties to devalue a class of people—in the usual example, women—even when that devaluation is not independently unlawful. The rules in place can of course be evaluated with reference to their contribution to inequality, but that is only to say that equality concerns are built into the rules in place. My concern is with the effect NIT should have on doctrine, taking as constant whatever array of values has shaped the rules in place.

37. The classic example is Sidis v. F.R. Publishing Corp., 113 F.2d 806 (2d Cir. 1940) (an action for invasion of privacy involving a magazine story depicting the plaintiff, who had been celebrated decades earlier as a child prodigy, as an eccentric, which harmed his reputation as well).

38. *See* Massaro & Norton, *supra* note 1, at 1654–83 (discussing this as a primary theme). *See also* Jeremy K. Kessler & David E. Pozen, *The Search for an Egalitarian First Amendment*, 118 COLUM. L. REV. 1953, 1959–60 (2018); Genevieve Lakier, *Imagining an Antisubordinating First Amendment*, 118 COLUM. L. REV. 2117, 2119–20 (2018) (discussing similar accounts). This reconsideration leads us to conclude that the rules in place were not on the Pareto frontier in the first place.

The conclusion of this argument, strikingly, is that the rules in place do not actually jointly optimize even in connection with the traditional media. Indeed, my strong impression is that advocates for new rules in light of NIT favored those rules before NIT became prominent, the most obvious example being advocates for the feminist argument for greater regulation of sexually explicit material.[39]

I stress that there is nothing wrong or even troubling about this kind of claim. Increasing the clarity with which we see how values are implicated by specific rules is a good thing, and if the rise of NIT generates clarity, all the better. It is just that the new rules, displacing the old ones, are not designed with any special characteristics of NIT in mind. The proposition that more harm is occurring is not built into the new rules in any way. That's not true of the second way that NIT might lead us to change the rules in place.

IV. HOW MORE SPEECH CAUSES MORE HARM, AND WHAT MIGHT BE DONE ABOUT IT

This Section examines several specific doctrinal areas and asks, how is more harm produced or distributed differently when this particular type of expression is disseminated through NIT rather than through the traditional media? If the rules in place strike the right balance between expression and harm, given the amount and distribution of harm that occurs when traditional media are used, they do so by affording constitutional protection to expression in terms set by the rules in place. NIT change the level and sometimes the distribution of harm, but we can maintain the same level and distribution of harm by protecting *less* expression than do the rules in place.[40]

A. The Implications of a Proportionality Rule in Place

Approaching the issue through the examination of selected doctrinal areas requires some defense. The reason is that doing so might seem unacceptably "categorical" to adherents of proportionality-based approaches to defining the constitutionally permissible limits of regulation.[41] Such approaches are common and perhaps even are predominant outside the United States.[42] Balancing approaches that used to be favored in the United States and that some

39. *See, e.g.*, CATHARINE A. MACKINNON, ONLY WORDS (1996).

40. We might think that the rules in place do not strike the right balance because we should put up with more harm than the rules in place allow. On some accounts that invoke institutional limitations on legislative and judicial competence, we actually think that the socially acceptable level of harm is higher than that generated by the rules in place. If so, the increased harm wrought by NIT might lead us to protect *more* expression so as to maintain the same, inadequate, level of harm.

41. For a prominent recent work advocating for such an approach, see JAMAL GREENE, HOW RIGHTS WENT WRONG: WHY OUR OBSESSION WITH RIGHTS IS TEARING AMERICA APART (2021).

42. For an overview of the doctrine and its use in many jurisdictions, *see* AHARON BARAK, PROPORTIONALITY: CONSTITUTIONAL RIGHTS AND THEIR LIMITATIONS (2012).

scholars continue to use are a form of proportionality as well.[43] Modern forms of proportionality review ask whether the legislature could reasonably believe that the regulation in question advanced some government interest other than mere suppression of expression, whether there were alternative responses to the utterance that would advance the government interest almost as effectively, and whether the impairment of free expression was proportionate to the harm the utterance caused.[44]

The large literature on proportionality suggests rather strongly that no one is advocating in favor of a case-specific proportionality rule.[45] Rather, they ask whether regulation of some category of expressions—such as hate speech or the Holocaust lie that Nazi Germany did not have a program aimed at the extermination of Europe's Jews—satisfies a structured proportionality test.[46] Put slightly differently, rather than examining proportionality with respect to a specific utterance, they examine the proportionality of regulation applied to categories of expression. These are categorical proportionality rules.

Now consider the effect of increasing the amount of expression subject to a categorical proportionality rule. It is easy enough to apply such a rule to expression via NIT: just plug the amount of harm into the final step and see whether the regulation is proportionate. One might find, for example, that false statements that injure reputation cause too little harm to justify the infringement of free expression when they are distributed in small newspapers and that the same statements distributed more widely multiply the harm sufficient to justify the infringement. Here, the same rule is applied—structured proportionality

---

43. *See* T. Alexander Aleinikoff, *Constitutional Law in the Age of Balancing*, 96 YALE L.J. 943 (1987).

44. The literature on proportionality in comparative constitutional law is quite large. *See, e.g.*, VICKI JACKSON & MARK TUSHNET, PROPORTIONALITY: NEW FRONTIERS, NEW CHALLENGES 1–2 (2017); GRANT HUSCROFT, BRADLEY W. MILLER, & GRÉGOIRE WEBBER, PROPORTIONALITY AND THE RULE OF LAW: RIGHTS, JUSTIFICATION, REASONING 1, 10–11, 16 (2014); AHARON BARAK, PROPORTIONALITY: CONSTITUTIONAL RIGHTS AND THEIR LIMITATIONS 1–2, 5 (2012).

45. The reason comes through in considering a standard example drawn from the German case law. The Basic Law recognizes a right to free development of the personality. A litigant claimed that this gave him the right to feed pigeons in a public square, contrary to a local ordinance. BVerfGE 54, 143. The court held that the right to development of personality was indeed infringed, but that the infringement was justified when the court applied the proportionality test. Now consider several variants. In one, the claimant is a socially isolated elderly person with few material resources whose sole pleasure in life comes from feeding the pigeons, while in another, the claimant is a well-to-do retired banker who simply enjoys feeding the pigeons. Social connectedness and wealth might well play some role in the proportionality inquiry. Contrast those cases with one in which the claimant wants to feed squirrels rather than pigeons. I doubt that proponents of proportionality review would contend that the animal species should play any role at all. The reason is that social connectedness and wealth are categorically relevant to the inquiry, whereas species membership is categorically not relevant.

46. *See, e.g.*, Vicki C. Jackson, *Constitutional Law in an Age of Proportionality*, 124 YALE L.J. 3094, 3111 n.79 (2015) (citing cases).

with respect to this category of statements—but produces different results for NIT.

We should pause here, though. The *form* of the argument is that application of the relevant proportionality rule should be sensitive to the amount of harm disseminating the statement causes and that that amount depends upon the medium through which the statement is transmitted. As the example just given suggests, that argument could (depending upon facts about different media) support different outcomes when a libelous statement is distributed by a small-circulation newspaper, a large-circulation one, radio, television, or another type of available media. Thus, I argue in connection with libel, nothing intrinsic to the structured proportionality approach rules out doing just that.[47]

What is more problematic is the thought that an utterance disseminated through *any* of the traditional media can be constitutionally protected but that the same utterance could be permissibly regulated when disseminated through NIT. What would it take to show, for example, that regulation of fake news distributed in radio broadcasts was disproportionate while the regulation of fake news distributed via NIT was proportionate? The answer, it seems, would have to be that the amount of harm in the latter case was large enough to tip the balance at the final stage of the proportionality inquiry. Add up all the harm caused by fake news distributed through the traditional media, and the regulation's infringement of free expression is too great, but add up all the harm caused by fake news over NIT, and its greater volume justifies the regulation's infringement. And that could be so even if the infringement in the NIT case was greater than that in the case of the traditional media if the increment of harm in the NIT case was disproportionate to the increment of infringement.

All this might be so, but it is unlikely that the facts will fall out in just the right way. In particular, I am skeptical that structured proportionality would draw a line between NIT and all traditional media rather than, for example, drawing the line between large-audience media—including some newspapers, and radio and television under some circumstances—and small-audience media. If that is correct, the effect of NIT is not to change the applicable rules, but, as discussed in Section IV above, to clarify the way the categorical proportionality rule in place should be applied.

So far, I have considered whether the increased amount of harm caused by expression via NIT could support a different rule for NIT and the traditional media. As others have observed, NIT makes new countermeasures available. As we will see, in several doctrinal areas, the ease with which a person or institution harmed by speech can counter that speech matters. Proportionality approaches explain why: Countermeasures are among the alternatives to regulation. In some settings, relying on the traditional media as the mechanism

47. Indeed, we can extract hints of such an approach from some libel and sedition cases. *See infra* Sections IV(B)–(C). I think it possible that we could do the same in other doctrinal areas, though I do not explore that possibility.

of response might be notably less effective than regulation, but in those same settings using NIT as a mechanism of response would be almost as effective as regulation.[48] Now the details.[49]

B. False Statements That Injure Reputation

Simply increasing the harm caused by false statements that injure reputation should not affect the damages rule once we know that the statement is not constitutionally protected. Suppose the liability rule is that actual damages can be awarded for false statements about public figures made with knowledge of their falsehood or with deliberate indifference to their truth or falsity—the *New York Times Co. v. Sullivan* liability rule.[50] The more widely such a statement is distributed, the greater the harm, and publication via NIT means that the statement will indeed be more widely distributed.

That argument is correct, but it looks at the problem in the wrong way. We should be concerned with whether NIT should lead us to reconsider the liability rule. Specifically, suppose a false statement injuring reputation is constitutionally protected when published in a newspaper because, for example, the plaintiff cannot show deliberate indifference. We should ask whether that same statement could lose its protection if disseminated via NIT. The answer, I believe, could be yes.

To see why we must ask why *any* false statement injuring reputation has constitutional protection. As Justice Powell famously observed, false statements injuring reputation have no social value in themselves.[51] We protect them nonetheless, out of a concern that imposing liability for false statements full stop would deter publishers from making some true statements out of concern that the institutions used to determine truth or falsity are imperfect. They sometimes may find false a statement that is actually true.[52] Publishers will then steer clear of the forbidden zone, as Justice Brennan put it.[53] The solution lies in devising a liability rule that produces the socially desirable amount of truthful statements and the socially acceptable level of harm to reputation.

Now consider how a person contemplating becoming a voluntary public figure would think: *I know that I will be exposed to the risk that my reputation*

---

48. For a discussion of this possibility, *see infra* Section IV(E).

49. The following subsections present the rationales for the rules in place dogmatically, without considering the merits of those rationales. Instead, I provide other scholarly references that support the dogmatic statements.

50. N.Y. Times Co. v. Sullivan, 376 U.S. 254, 279–80 (1964).

51. Gertz v. Robert Welch, Inc., 418 U.S. 323, 340 (1974).

52. Note that publishers are indifferent about the possibility that the decision-makers will mistakenly determine that a false statement is actually true because such a finding would let them off the hook entirely.

53. Speiser v. Randall, 357 U.S. 513, 526 (1958). This is the mechanism of the well-known chilling effect. *See* Frederick Schauer, *Fear, Risk, and the First Amendment: Unraveling the Chilling Effect*, 58 BOS. U. L. REV. 685, 688–89 (1978) (providing a detailed analysis).

will be injured by false statements, but the benefits I'll get from being a public figure outweigh those risks. The liability rule in place is set by a judgment that it gets us enough people willing to be public figures.[54] Now increase the amount of injury to which the potential public figure is exposed. Someone who would choose to become a public figure given some risk of harm to reputation might choose not to if the risk of harm increases.[55] And, in turn, that means that the pool of potential public figures shrinks below the level we thought we needed. To bring it back to that level, we have to adjust the liability rule so that some statements that were protected under the liability rule in place are not protected under the new rule.

It is obvious how NIT fit into this picture. They increase the amount of harm, and under the liability rule in place, more people will choose not to become public figures. We must adjust the liability rule to keep the pool of potential public figures where we want it to be.

As noted earlier, we could do that in one of several ways. We could retain the liability rule in place for libelous statements distributed through traditional media but create a looser rule for libelous statements distributed via NIT. Or we could change the rule to the looser one and apply it to all media.

One suggestion that rattles around in the literature on libel and in the comparative case law hints at the possibility of something like the first approach—one that differentiates between NIT and the traditional media. That suggestion has one primary component and one secondary component. The primary component is that libelous statements should be constitutionally protected only if they were made after conforming to the standards of responsible journalism.[56] The secondary component is that those standards differ depending on circumstances. The obvious example is that responsible journalism might require more extended inquiries for stories that the journalist works on for a long time than are possible in connection with breaking news. In one celebrated case, the court considered how much effort a responsible journalist would expend in trying to get a comment from the libeled person.[57]

This second component might support different rules for small-circulation publications and large-circulation ones or for citizen journalists and professional reporters. In the second of each pair, the available resources might be larger, and so responsible journalism might require more effort. Adapting this

---

54. This is the specific form that the question of joint optimization takes here.

55. It is easier to see this phenomenon if we think about people deciding whether to run for public office and become public figures, but I think the logic operates in the same way when people decide to embark upon careers that lead them to become celebrities or, as in one important case, football coaches. *See* Curtis Publ'g Co. v. Butts, 388 U.S. 130 (1967). In traditional law-and-economics terms, this is the effect a rule has on the activity level.

56. *See* Grant v. Torstar Corp., [2009] 3 S.C.R. 640, 642 (Can.) (adopting a defense "that would allow publishers to escape liability if they can establish that they acted responsibly in attempting to verify the information on a matter of public interest").

57. Jameel v. Wall St. J. [2006] UKHL 44, [2007] 1 AC 359, [53], [54], [58] (HL).

argument to our context, we might then end up with one liability rule for libelous statements in the traditional media and another for the same statements made via NIT.

C. Expression that Induces Law Violation

The classic law of sedition involves prosecution for statements criticizing existing government policy, typically where the policy is not itself unconstitutional.[58] The prosecutions do not rest on the proposition that criticism is unlawful in itself. Rather, the government claims that criticism of its lawful policies increases the risk that someone will act unlawfully, interfere with the military draft, or engage in violent actions aimed at overthrowing the government that generated the policy being criticized.[59]

Criticism has these effects because it can lead listeners to conclude, on their own, that something should be done about the policy other than seeking its repeal by lawful means, as compared with seditious speech that persuades listeners to act. Criticism can take varying forms: Speakers can drily offer reasons for rejecting the policy, package those reasons in rhetorical forms that make them more effective, and urge listeners to do what they can to undermine the policy. This is further complicated because there is inevitably a temporal gap between the speaker's utterance and the listeners' unlawful action. That generates the principle that the remedy for bad speech is more speech.[60] That is, the speaker offers reasons for rejecting the policy by acting unlawfully. The government and its supporters can prevent the unlawful action by offering their own reasons for the policy or for attempting to change the policy by repealing it.

The law of free expression recognizes that sometimes the temporal gap is too small for counter-speech to be effective. The time between the speech and the action might be quite short, as in a variant of the well-known case of *Hess v. Indiana*.[61] A speaker before an angry crowd says, "[w]e'll take the fucking street again."[62] Under the circumstances, the possibility that counter-speech will be effective is tiny. The gap can be small in another sense. The speaker's rhetoric may be so powerful that it blocks the effectiveness of counter-speech even when there is a larger temporal gap between the speech and the ensuing action.

---

58. *See, e.g.*, Debs v.United States, 249 U.S. 211 (1919) (upholding a conviction for obstructing the military draft by giving a speech vividly asserting that the draft was unjust).

59. *See, e.g.*, Schenck v. United States, 249 U.S. 47 (1919) (upholding a conviction for obstructing the draft, arguing that the defendant's speech encouraged listeners to resist the draft); Dennis v. United States, 341 U.S. 494 (1951) (upholding convictions for conspiring to overthrow the United States government by force and violence).

60. In the proportionality analyses, counter speech is one of the available alternatives to regulation.

61. Hess v. Indiana, 414 U.S. 105, 108 (1973) (quoting Hess v. State, 297 N.E.2d 413, 415 (Ind. 1973)).

62. The statement in *Hess* was, "'[w]e'll take the fucking street later,' or '[w]e'll take the fucking street again.'" *Id.* at 107.

Finally, the law of free expression can focus on personal responsibility. Listeners who actually violate the substantive law can, of course, be prosecuted for doing so without giving any consideration to whether punishing the speech would be constitutional.

These considerations generate the rules in place dealing with liability for disseminating words that cause unlawful action or, more precisely, that increase the risk that unlawful action will occur. Those rules say something about how large the temporal gap is—"clear and present" in some formulations, "imminent" in others—the words' rhetorical form—"incitement" or "mere criticism," for example—and the preference for imposing liability on those who violate the substantive law.[63] My concern here is not with what the right rule in place should be. Instead, it is whether the increase in expression occasioned by the development of NIT should change whatever the rule in place is.

It is difficult to see how that increase might affect the way in which the rules address either the temporal gap or the question of rhetorical effect. Expression via NIT is almost certainly as far removed in time from the listeners' unlawful action as it is in many traditional-media cases.[64] Any rule that effectively deals with the temporal gap for expression via the traditional media would likely work for expression via NIT. Nor does it seem likely that the rhetorical effects differ much. It is difficult to see how a statement broadcast over the radio would be less effective than one distributed via NIT. The non-cognitive effects associated with the traditional media, which might include things like the cinematographic effects in Leni Riefenstahl's *Triumph of the Will*, have their counterparts in NIT.[65] There are not any obvious reasons to think that the response of the rules

63. *See* Dennis v. United States, 341 U.S. 494 (1951) (applying a "clear and present danger" test); Brandenburg v. Ohio, 395 U.S. 444 (1969) (applying a test requiring imminence and incitement). There are well-known problems with essentially every formulation of the candidate rules in place: the clever inciter (Marc Antony at Caesar's funeral is the usual example, as in Susan Benesch, *Inciting Genocide, Pleading Free Speech*, 21 WORLD POLICY J. 62, 67 (2004)), the sophisticated publisher whose arguments sway a simple-minded soldier (Abraham Lincoln's concern, in a letter to Erastus Corning, June 12, 1863, available at http://www.abrahamlincoln online.org/lincoln/speeches/corning.htm), or the speaker who offers dry reasons in a setting where listeners are likely to act unlawfully quite quickly (John Stuart Mill's example of the speaker who tells an excited crowd that corn dealers are oppressors of the poor, in JOHN STUART MILL, ON LIBERTY, ch. 3 (1859)). The law of free expression is best understood as holding that these problems are worth tolerating because the rule in place optimizes speech and harm even if the clever inciter, the publisher, or the dry speaker escape liability that in principle they should bear. The thought is that any rule that captured these anomalous cases would be suboptimal because of the alternative rule's effects on more standard cases.

64. Massaro and Norton assert, "Time for counterspeech is now vanishingly small." Massaro & Norton, *supra* note 1, at 1645. They have in mind speech such as threats, which cause harm immediately upon receipt, rather than speech that increases the risk of law-breaking. *Id.* Yet, there is no time for counterspeech when true threats are made via traditional media.

65. For a description of these effects, see "Hand-Held History," The Economist, Sept. 3, 2003 (obituary).

in place to such rhetorical effects in the traditional media would be less adequate for NIT.

What the increase in expression might do, though, is increase the audience. More people might "hear" expression via NIT than hear it through in-person speeches, newspapers, or television. For ease of exposition, I will assume that a specific expression we are concerned with leads a constant percentage of the audience to break the law.[66] Even if that percentage is quite small, increasing the audience would increase the number of unlawful acts.

And that matters for the joint optimization calculation, taking into account the preference for imposing liability on the law-breakers when possible. The thought is this: where some expression induces only a handful of people to break the law, and sufficient enforcement resources can be devoted to prosecuting law-breakers to push down the harm actually inflicted. Liability is imposed on the speaker as a backup to push that level down to where society wants it to be. As the number of law-breakers increases, the enforcement resources devoted to catching them might increase. At some point, though, enough resources cannot be devoted to enforcement against the law-breakers to push the level down as much without affecting the ability to do other things, which include enforcing other laws, providing education, providing medical care, and so on through the public budget.[67]

The conclusion then is obvious. To maintain the jointly optimized level, it should be easier to hold liable speakers who send their messages via NIT. This might be a reason for having two rules, one that makes it difficult to hold someone liable for expression disseminated via traditional media and one that makes it easier to do so where exactly the same message is disseminated via NIT.[68]

D. Sexually Explicit Material

In the pre-Internet world, a person who wanted access to pornography or obscenity would have to go somewhere to get it, such as an adult bookstore, a theater showing pornographic films, or a store selling or renting videotapes and CDs. Today, that same material is available at home and on-demand. Producing sexually explicit material for distribution via the Internet is so much cheaper

66. I cannot think of a reason why dissemination via NIT would *decrease* the percentage of listeners who are induced to break the law, though I cannot rule out the possibility that someone with more imagination than I could come up with some such reason.

67. With some effort, we can find hints of this thought in Justice Holmes's reference to the "puny anonymities" in *Abrams v. United States*, whose speech he would have held free from punishment. Abrams v. United States, 250 U.S. 616, 629, 631 (1919) (Holmes, J., dissenting). In contrast to his willingness to uphold the conviction of Eugene V. Debs, a prominent political figure and eloquent speaker. Debs v. United States, 249 U.S. 211, 216 (1919).

68. To use the example of U.S. law, the United States might continue to use the strict *Brandenburg* "incitement to imminent lawless action" test for expression via the traditional media but revert to a looser "clear and present danger" test for expression via NIT. Brandenburg v. Ohio, 395 U.S. 444, 448–49 (1969). *Cf. id.* at 449 (Black, J., concurring).

than producing films that an extremely large supply of amateur pornography is now available.[69] NIT, in short, rather clearly did increase the amount of sexually explicit material available to consumers. If such material causes harm, as those who wish to regulate it believe, NIT have increased the amount of harm. How might that matter for the structure of free expression law in this setting?[70]

First, the harms said to be caused by the distribution of sexually explicit material and the mechanisms by which the harms occur must be identified. The standard arguments are that there are two types of harms: an increased risk of assaults upon women—or those sometimes described as "in the position of women," specifically LGBTQI individuals—and a diminution in women's social standing and credibility.[71] The primary causal mechanism, and the one on which I focus, is non-cognitive: the unconscious effect that viewing sexually explicit material has on the viewer.[72] Further, some treatments of sexual matters are unquestionably protected by free expression principles. For example, discussions of the value of adultery, even when they are presented in ways that have some non-cognitive components, should be protected by the First Amendment.[73]

Consider first the claim about sexual assault that the distribution of the material in question causes an increased risk of ordinary law-breaking. As I developed earlier, the increased risk of law-breaking caused by the proliferation of the material at issue might justify modifying the rule in place if the other resources for deterring law violation cannot be increased without diminishing social welfare along constitutionally relevant dimensions.[74]

There is one wrinkle. The prior analysis of this issue considered the non-cognitive features of expression said to cause an increased risk of law-breaking through its acknowledgment that effective rhetoric might enhance the cognitive arguments for the targeted action. There the non-cognitive features were secondary to the cognitive ones. Here they are primary. Still, the prior analysis suggested the possibility that the rules in place might be altered, or at least

---

69. For a discussion of the comparative economics of producing sexually explicit material for videos and the internet, *see* Cade Metz, "The Porn Business Isn't Anything Like You Think It Is," Wired Magazine, Oct. 15, 2015, available at https://www.wired.com/2015/10/the-porn-business-isnt-anything-like-you-think-it-is/.

70. The proliferation of sexually explicit material might lead us to reconsider the balance of values the rules in place strike, but I have nothing to add here to the analysis of that proposition offered in Section III, *supra*.

71. I include "credibility" in this formulation to capture the idea often describing as the "silencing" effect of sexually explicit material. That effect is sometimes real but more often metaphorical—it is not primarily that the distribution of sexually explicitly material causes women to stop offering their views, though that can occur, but rather that it leads to an undervaluation of those views' merits.

72. *See, e.g.*, MACKINNON, *supra* note 39.

73. The example of course is of D.H. Lawrence's book, *Lady Chatterley's Lover*. D.H. LAWRENCE, LADY CHATTERLEY'S LOVER (1928).

74. *See supra* Section IV(C).

applied differently, as the non-cognitive features of the expression increased because the required causal connection between the expression and actual law violation might be weakened. So too here. Perhaps the increased risk of sexual assault occasioned by an increase in the distribution of sexually explicit material via NIT could justify weakening either the rule in place or its application.[75] But all this would be triggered only if we thought that the increased risks of sexual assault rendered other modes of deterring such assaults inadequate.

Now to the claim about the effects of sexually explicit material on women's social standing and credibility. Here the causal mechanism is entirely non-cognitive. That matters because the usual response to harmful speech—more speech—is basically unavailable (because the counter speech is typically understood to be primarily cognitive, with the possibility that effective rhetoric might somewhat enhance the counter-speech's effect).

The rule in place can take one of several approaches to primarily non-cognitive expression. First, the rule might simply disregard the non-cognitive causal mechanism. Where it does, increasing the amount of material that works through such a mechanism would not justify altering the rule in place.

Other approaches end up at the same point. The rule might treat the non-cognitive mechanism in the same way it treats the cognitive one. Or, we might have built non-cognitive features of expression into the rule in place because of a combination of "family resemblance" and institutional concerns. With two co-authors I have argued that this is how U.S. free expression law deals with what we call "free speech beyond words": nonrepresentational art, instrumental music, and literal nonsense.[76] We have one rule for expression that works through cognitive mechanisms. We apply the same rule to expression whose effects work through non-cognitive mechanisms but not because there is a consistent or reasoned principle. We do so because those expressions are enough like a cognitive expression that decision-makers, especially courts, would find it difficult to distinguish the two types of speech effectively because they could not reliably say that in one case, the non-cognitive features were merely secondary and in another say that they were primary. If we take either of these approaches, the analysis in Section V(C) kicks in.

Finally, and most distinctive, the rule in place may, or may not, take non-cognitive effects fully into account. For example, the rule might be that the effects upon women's social standing and credibility do occur but are not widespread enough to justify some regulation of sexually explicit material in light of the adverse effect regulation might have on the distribution of unquestionably protected treatments of sexual matters. The rule in place denies

---

75. As before, this could occur either by shifting from something like a *Brandenburg* rule requiring imminent law-breaking to the kind of balancing called for by a "clear and present danger" approach, or by adjusting the weight given to the danger within such a balancing approach. Brandenburg, 395 U.S. at 449 (Black, J., concurring).

76. *See generally* MARK TUSHNET, ALAN CHEN & JOSEPH BLOCHER, FREE SPEECH BEYOND WORDS: THE SURPRISING REACH OF THE FIRST AMENDMENT (2017).

the government the power to regulate sexually explicit meat-world photographs or films. The increased amount of material distributed via NIT might be enough to push the harm across the threshold. The NIT-responsive rule might expand the definition of unprotected material, for example, from obscenity to a broader category of pornography, narrowing the class of protected material and thereby pushing the amount of harm caused by such material below the threshold.

What if the rule in place already allows regulation of sexually explicit material because the harm it causes is large enough? Here too, the analysis should be simple. The more such material is available, the greater the harm. As before, the rule in place optimizes expression and harm, and so it should be adjusted when harm increases. As I argued in Section V(C), the adjustment, where the harm is the risk of law-breaking, ordinarily would take the form of loosening the degree of a causal connection between expression and law-breaking. That adjustment is not available here because the non-cognitive effects occur immediately upon exposure to the material. So, as suggested before, the adjustment here probably should take the form of expanding the definition of unprotected sexually explicit material.

E. False Statements That Cause Systemic but Not Material Harm: Herein of Fake News

The Stolen Valor Act made it unlawful for a person to claim falsely that he or she had received certain military honors. The U.S. Supreme Court held the Act unconstitutional in *United States v. Alvarez*.[77] Justice Kennedy's plurality opinion characterized the statute as criminalizing a false statement that caused no material harm in contrast to statutes criminalizing résumé fraud or, had there been such a statute, one criminalizing falsely claiming to have received military honors for the purpose of obtaining leniency in the criminal justice system or other material benefits.[78] Were the government to have the power to do that, Justice Kennedy argued, it would have the power to criminalize a wide range of false statements, some of which—such as "social lies"—might have some social value and others of which caused no harm at all.[79]

Because the Stolen Valor Act was not a general ban on false statements, the Court had to explain why the government could not have the power to identify a specific category of false statements that did not cause material harm but might cause other harms.[80] The government defended the Act in part on the ground that false claims about military honors undermined the system of military honors by creating a world in which such honors received less social support.[81] Justice

77. United States v. Alvarez, 567 U.S. 709, 715 (2012).

78. *Id.* at 720–22.

79. *Id.* at 748–79. Justice Breyer's opinion concurring in the judgment refers explicitly to false statements that "can serve useful human objectives." *Id.* at 733 (Breyer, J., concurring).

80. *Id.* at 720–22 (plurality opinion).

81. *Id.* at 716.

Kennedy was skeptical about the claim that there would be such systemic harm but ultimately accepted its possibility.[82] On that assumption, the Court offered one response and hinted at another. Whatever adverse systemic effect occurred, the Court said, could be overcome by a form of counter-speech—maintaining an easily accessible location (a website, specifically) that could be checked to determine whether a person actually had received the honor she or he claimed.[83] Further, the Court suggested that systemic harms, while real, could not be addressed by the government through regulations that infringed free expression rights.[84] Avoiding systemic harms might not be a permissible government purpose for free expression purposes.

The Court raised additional concerns. Acknowledging a power in the government to identify and punish falsehoods raised the specter of a Ministry of Truth, authorized to identify government-approved truths and disapproved falsehoods, and (implicitly) we cannot trust legislators to exercise that authority in a politically neutral way, even if their decisions were subject to judicial review.[85] Other constitutional systems are less skeptical about legislative and judicial capacity, allowing legislators to outlaw claims that the Holocaust, appropriately defined, did not occur.[86]

The Stolen Valor Act decision and the Holocaust lie case are precursors to today's concern about fake news. Those concerned about fake news typically point to systemic rather than material harms—such as undermining the people's confidence in public institutions. Those cases also identify some specific forms of fake news that might be thought to cause material harm. The publication of false information about where, when, and how to vote by people who know that information is false is a good example.

Suppose a statement causing systemic but not material harm is made via NIT rather than via the traditional media. How might the constitutional analysis differ in light of the larger harm inflicted?

Three considerations count against changing the analysis. If only material harm counts for free expression law, and systemic harm does not, increasing the amount of systemic harm should not change anything. Skepticism about legislative and judicial capacity seems unrelated to the medium through which expression occurs. If we do not think that the Ministry of Truth can reliably determine what is true when published in a newspaper, we should not think that it can do a better job when statements are distributed via NIT. In addition,

---

82. *Id.* at 726.

83. *Id.* at 729.

84. *Id.* at 725 (requiring a "direct causal link" between the regulated expression and the occurrence of systemic harm).

85. *Id.* at 723. This concern would seem to be fully addressed by a requirement that people could be punished for uttering specific government-identified falsehoods only if the speaker knew that their statements were false.

86. BVerfGE, 1 BvR 23/94, Apr. 13, 1994, https://law.utexas.edu/transnational/foreign-law-translations/german/case.php?id=621 (the "Auschwitz lie" decision).

Alvarez might rest on the assumption that, with respect to claims about having received military honors, the systemic harm not countered by reasonably available countermeasures was below some threshold. If the increased systemic harm exceeds that threshold, the rule in place might itself allow the imposition of liability—and so the rule would not need adjusting when dissemination of material via NIT also exceeds that threshold.[87]

That leaves a final consideration that might count in favor of changing the analysis. To the extent that countermeasures such as the online list of recipients of military honors matter, the increased systemic harm might make a difference. New technologies make new modes of response available, of course, as the online list example shows, but the rate at which systemic harm increases might outpace the rate at which new technologies make countermeasures available.[88] If so, the liability rule might be adjusted to take account of the comparative ineffectiveness of countermeasures.

Clearly, then, the implications of NIT depend crucially upon what the rules in place are. The *Alvarez* decision suggests that the rules in the United States would not allow a ban on fake news either via traditional media or NIT.[89] In contrast, a law banning the distribution of false information about voting, knowing the information to be false, might be constitutionally permissible because such information causes material harm. For example, votes may become actually lost when voters think that an election is on a Wednesday rather than on a Tuesday.

In contrast, the rules in place that constitutionally permit a ban on the Holocaust lie rest on two predicates different from those used in U.S. law.[90] Systemic harm clearly counts, and there is less skepticism about legislative and judicial capacity.[91] Finally, the fake news issue seems to be almost a paradigm of a problem where imposing liability for the systemic harm caused by fake news disseminated via the traditional media might be disproportionate. Conversely, doing so for material disseminated via NIT would satisfy a proportionality requirement because of the greater systemic harm.

---

87. *But cf.* Massaro & Norton, *supra* note 1, at 1643 ("Falsehoods that allege widespread voter fraud . . . corrosive[ly] fuel[] voter cynicism and disengagement."); Stone, *supra* note 17, at 47 ("[O]ne might readily argue that widespread lies that are intended to distort public discourse rise to . . . [a significant] level of harm" akin to the harm associated with "defamation, perjury, and fraud."). The analogy here would be to impose liability for making statements that a reasonable person would understand to be threats of harm. *See* Section IV(F) *infra*.

88. I cannot vouch for the following example from my own knowledge, but I have been told that people creating "deep fakes" and those developing techniques to detect such fakes are in an arms race that the former are winning.

89. *Alvarez*, 567 U.S. at 723. Recall that if *Alvarez* adopts a threshold requirement to taking systemic harm into account, and the harm of fake news exceeds that threshold, we do not have to change the rule to allow regulation of fake news.

90. BVerfGE, 1 BvR 23/94, Apr. 13, 1994, https://law.utexas.edu/transnational/foreign-law-translations/german/case.php?id=621.

91. *Id.*

F. Threats: Herein of Material Harm, Distress and Anxiety, and Cyberstalking

Threats cause two kinds of harm. Their targets may experience distress and anxiety upon receiving the threat, and they may change their behavior based on a threat by refraining from going out in public as often as before, purchasing a home security system, and take other precautions.[92] Under current doctrine regulating liability for threats, the government can prohibit "true threats" delivered in person, by mail, or through NIT. Here, joint optimization takes the form of defining what counts as a true threat. Importantly, the definition assumes that people in a complex society have to put up with some unwanted experiences that cause them distress or anxiety. This principle is embodied in the idea that the government cannot punish speech if the recipient could reduce its impact by "merely" averting her eyes.[93] Distress or anxiety occurs immediately upon receiving the distressing material and averting one's eyes is thought to cap the amount of distress or anxiety. By this logic, distress or anxiety below some threshold does not count as harm.

If pressed, courts might recognize that the material harm caused by threats—changes in behavior—also must cross a threshold. So, for example, changing one's voicemail message might not count, but purchasing a home security system might. More important, implicit in the "avert your eyes" principle is the thought that some people might venture into the domain of expression only to withdraw when their interventions elicit responses that offend them. Such withdrawals are changes in behavior, which I have characterized as material harm. The "avert your eyes" principle might assume that only a tolerably small number of people will do so. But, if the number increases because interventions elicit dramatically more offensive responses, we might conclude that a threshold has been crossed.

Under current law, a true threat consists of words (or images) that the target would take to be a *real* threat—that is, one with some likelihood of actually being carried out.[94] This might be a subjective standard: did the recipient actually understand the words to be a real threat? It might be objective: would a reasonable recipient understand the words to be a real threat? A similar mental state issue arises with respect to the person making the threat: can he be held

92. A low-level example: When I received something that was on the borderline between hyperbolic criticism and a threat, I changed my telephone voicemail response to eliminate my home telephone number.

93. Erzoznik v. Jacksonville, 422 U.S. 205, 210 (1975) (holding unconstitutional an ordinance prohibiting the showing of nudity in films at drive-in movie theaters, where the scenes could be viewed by people not attending the showing: "[m]uch that we encounter offends our esthetic, if not our political and moral, sensibilities. Nevertheless, the Constitution does not permit government to decide which types of otherwise protected speech are sufficiently offensive to require protection for the unwilling listener or viewer").

94. *See* Elonis v. United States, 575 U.S. 723, 737–38 (2015).

liable only if he specifically intends the words to be understood as a real threat?[95] Can he be held liable if, with reckless disregard of his words' effect, he utters words that the target—or a reasonable person—would take to be a real threat? Can he be held liable if the target takes the words to be a real threat even if he did not intend them to be so? Current doctrine is somewhat unsettled on these questions.[96]

For present purposes, the question at hand is whether or not to alter the definition of true threats because of NIT—either for all threats or only for threats delivered via NIT. Consider first the threshold level of distress and anxiety. At first glance, this would appear to operate at the individual level. Each of us has to put up with distress or anxiety below the threshold, as a cost of living in a "pluralistic society, constantly proliferating new and ingenious forms of expression."[97] Yet, if threats proliferate (or get distributed differently), we might conclude that the aggregate level of distress and anxiety is too high even if no individual experiences any more distress or anxiety than before. If so, it may be prudent to lower the threshold that is built into the definition of a true threat.

Next, consider words conveyed via a traditional medium such as a letter or a television address in which the speaker specifically identifies a target. One might conclude that those words are not a true threat. However, when conveyed via some NIT, the same words might be a true threat because the medium affects the seriousness with which the words are reasonably taken. Note that here, NIT does not change the rule in place—liability for true threats—but it might change the application of the rule to identical words.

Threat liability brings to the fore mental state issues that arise in connection with related activities such as cyberstalking and revenge porn.[98] These activities cause the same harms that threats do, material harm by inducing the target to change her behavior, distress, and anxiety. Consider first cyberstalking. Appropriately defined to mean something like lurking in a person's physical presence with the intent of causing distress and anxiety, stalking in the so-called "meat world" can be criminalized or made the basis of civil liability.[99] Suppose

---

95. I should be explicit here. I think that the First Amendment is no barrier to imposing liability for words that were specifically intended to be a real threat and that would reasonably be understood as a real threat.

96. *Elonis*, 575 U.S. at 726 (holding as a matter of statutory interpretation that the federal threat statute required the government to show that a reasonable person would take the words in question as a threat); *see also* State v. Boettger, 450 P.3d 805, 817 (Kan. 2019), *cert. denied*, 2020 U.S. LEXIS 3370 (U.S. June 22, 2020) (denying certiorari to the Court, in 2020, to review to a decision holding that the First Amendment barred the government from imposing threat liability where the defendant acted with reckless disregard of the words' impact on the target).

97. *Erzoznik*, 422 U.S. at 210.

98. *See* MARY ANNE FRANKS, THE CULT OF THE CONSTITUTION, 45–47 (2019); *see also* DANIELLE K. CITRON, HATE CRIMES IN CYBERSPACE 187–88 (2014).

99. For an overview of the requirements for a conviction for stalking under federal law, *see* Deborah C. England, "Federal Stalking and Harassment Laws," CriminalDefenseLawyer, available at https://www.criminaldefenselawyer.com/resources/federal-stalking-and-harassment-laws.htm.

an individual cyberstalks a victim by sending the victim repeated messages that cause anxiety and distress above the threshold required for threat liability—even if the messages are not true threats. It is hard to see why the theory—according to which liability for true threats is not prohibited by free expression—would not be applicable to this cyberstalker.[100]

Now, suppose that cyberstalking is defined to include uncoordinated actions by individuals, each of whom sends one or two messages that do not cross the required threshold of distress and anxiety. In the aggregate, these hundreds of messages do induce the target to change her behavior and inflict more than the required level of distress and anxiety. At this point, details matter. Is the proposal to make each participant jointly and severally liable for all of the harm? For a proportionate share of the harm? Importantly for liability to attach, must the participant know that others are sending harassing messages; or would reckless disregard of that possibility be enough? Alternatively, could it be an unreasonable failure to consider that possibility? The answers to these questions will determine how much imposing liability will deter people from sending distasteful but constitutionally protected messages—that is, it will determine the chilling effect of liability on expression. That, in turn, will affect how many people withdraw from participating in a public discussion because they are unwilling to put up with the attendant harassment. Compare an actual knowledge standard with a reckless one—the former being more difficult to prove, there will be a smaller chilling effect but a larger number of people withdrawing from participation. This is once again the joint optimization question. I defer consideration of these issues to Section VI(C), where they arise in connection with more realistic proposals to make platforms liable for distributing the messages that, in the aggregate, are cyberstalking.

This is similar for revenge porn. It is clear that the privacy considerations that shape the free expression rules in place would allow a state to impose criminal or civil liability on a person who distributes, hand-to-hand, sexually explicit photographs taken of another when he knows that the person photographed agreed to be photographed only on condition that the material would not be distributed at all.[101] Suppose the target can prove only that a reasonable person who knew the circumstances under which the photograph was taken would have inferred that she agreed to be photographed only on that same condition. Further, suppose that the material is distributed (again, hand-to-hand) by someone other than the photographer. Can liability be imposed for doing so if a reasonable person seeing the photograph would have inferred that it was taken with the same restrictive condition on distribution? Here too, the mental element

100. *See* United States v. Stevens, 559 U.S. 460, 472 (2010).

101. The analogy would be to the tort of intentional infliction of emotional distress (IIED). The Court limited the scope of IIED liability when the speaker intentionally inflicted emotional distress in the course of saying something about a matter of *public* interest. *See* Snyder v. Phelps, 562 U.S. 443, 452 (2011). The holding in this case casts no doubt upon the tort's purchase in connection with purely private matters, such as arise in connection with revenge porn.

question leads one to consider the chilling effect. But note that these questions arise when considering liability for hand-to-hand distribution. It is not at all obvious that the answers would be different if the distribution occurred via NIT—at least when we deal with liability on the photographer or individual distributor. And again, I defer consideration of the chilling effect issue when liability is imposed on the platform.

United States v. Stevens might be thought to cast doubt on these conclusions; here, the Court expressed its unwillingness to identify what it called "new categories of speech outside the scope of the First Amendment."[102] The focus, according to the Court, should be on categories of speech that have been "historically unprotected."[103] These statements might be taken to weigh against the constitutionality of laws against revenge porn and cyberstalking, neither of which, it might be said, were historically unprotected categories.

Yet, the Court acknowledged at least the possibility that there might be some such categories that "have not yet been specifically identified or discussed as such in our case law."[104]

And it observed that its reluctance to create a new category was tied to the government's argument for what the Court called "a simple cost-benefit analysis."[105] Perhaps *other* arguments might support creating new categories or, perhaps better, support a new understanding of the scope of the historically unprotected categories. The argument that cyberstalking could not have been historically unprotected because it did not exist in 1791 is, of course, frivolous. On the other hand, the argument that cyberstalking is sufficiently analogous to a threat, and thus within a historically recognized unprotected category, is not.

Nor does the analogical argument rely upon a simple cost-benefit analysis. The structure of that argument is this: true threats were historically unprotected because they inflicted material harm and caused distress and anxiety, in both instances above some socially acceptable threshold. Revenge porn and cyberstalking do the same thing and so fall within the historically recognized category loosely but inaccurately described as true threats. Whether one finds the analogy compelling is another matter. *Stevens* does not rule out the possibility that the analogical argument is sufficient.

In sum, the First Amendment rules in place regarding threat liability, especially their mental state components, work just as well for NIT as for traditional media. Free expression principles should not be a barrier to laws against revenge porn or cyberstalking. And, if those phenomena are a social

---

102. *Stevens*, 559 U.S. at 472.

103. *Id.*

104. *Id.*

105. *Id.* at 471. This is an appropriate place to note that constitutional systems relying on structured proportionality would have no difficulty in doing a cost-benefit analysis and so would have no difficulty accommodating laws against cyberstalking and revenge porn if, of course, such laws were thought to conform to the socially acceptable balance between speech and harm.

problem, then the legal problem is legislative failure to act, not freedom of expression.

V. REGULATING PLATFORMS

By definition, NIT are new as they distribute material using technologies different from those used to distribute material via traditional media. This is especially true when NIT are in private hands as they give rise to seemingly new business models that are perhaps not merely *newspapers on a really large scale*. Or so some advocates for the regulation of Twitter, Facebook, Google, and other platforms for distributing material argue.

The idea of platform regulation has generated a large amount of literature, and I address only the narrow question of what the free expression rules in place have to say about several forms of platform regulation. As throughout this Essay, the underlying question is whether the platforms' business models and operations are such as to require adjustment to the rules in place so that we can maintain the balance between speech and harm that the rules strike.

I consider three forms of platform regulation: platforms as a form of government; application of general laws to platforms; and platform liability for transmitting expressions by others where those expressions are themselves unprotected by principles of free expression.

A. Platforms as Governments

The Court in *Packingham v. North Carolina* described "cyberspace" and "social media" as among "the most important places . . . for the exchange of views," in "the modern public square."[106] Similar observations have generated arguments that platforms, though privately-owned, should be treated as state actors.[107] As such, platforms could not discriminate among users on the basis of the content the users either posted to the platform or downloaded from it. Nor could they bar users from posting material that was protected by the Constitution, though they could choose to allow the dissemination of unprotected material just as a government can refrain from exercising the regulatory powers it has.[108]

Nearly all the Supreme Court's state-action cases over the past several decades, including cases invoking the First Amendment, have refused to treat private parties as state actors.[109] For present purposes, all that needs to be said is that it is not clear why the free expression aspects of NIT, including the

106. *See* Packingham v. North Carolina, 137 S. Ct. 1730, 1735–37 (2017).

107. *See, e.g.*, Wu, *supra* note 1, at 37 (discussing but rejecting the argument).

108. If platforms allow their use to disseminate unprotected material, they might be liable as accomplices, a possibility addressed in Section V(C) *infra*.

109. *See, e.g.*, Manhattan Cmty. Access Corp. v. Halleck, 139 S. Ct. 1921, 1934 (2019) (holding that the operator of a public access channel on a cable television system was not a state actor).

business models they enable, should *change* the state-action doctrine. The messy and confusing doctrine already has within it the resources to support both the conclusion that platforms are state actors and the conclusion that they are not.

B. Platforms and General Laws

Newspaper and television broadcasters have to comply with labor laws, provide safe workplaces as required by occupational safety and health laws, and honor their contractual obligations.[110] These are, as the Supreme Court has put it, "generally applicable laws [that] do not offend the First Amendment simply because their enforcement against the press has incidental effects on its ability to gather and report the news."[111] In contrast, laws that single out the media for regulation either expressly or by gerrymandering seemingly general criteria so that only the media, or some subset of the media, fall within the regulation's scope, have to survive either intermediate or strict scrutiny.[112]

The first of these principles support the application of other general laws to NIT platforms. Such laws include antitrust and competition law, fiduciary law,[113] and the law of common carriers. For example, if platforms fit within the general category of common carriers, they must treat all users equally. This nondiscrimination principle might produce net neutrality.

The key problem is that the applicable laws must be truly general; they cannot single out information media or, even more, a shortlist of specific platforms for regulation. And applying antitrust, fiduciary, or common carrier laws to platforms might do just that—though it might not if the application requires too much tinkering with existing law.

To ease exposition, I focus on fiduciary law, though the difficulties can arise with respect to antitrust and common carrier law. Skeptics suggest that platforms do not satisfy existing criteria for identifying entities with fiduciary obligations.[114] At first look, treating platforms as fiduciaries would be to treat them specially, not generally; however, that is not quite right as fiduciary law evolves incrementally. Presented with arguments that some entity previously outside the scope of fiduciary law be brought within it, courts ask not only

---

110. *See, e.g.*, Assoc. Press v. NLRB, 301, U.S. 103, 132–33 (1937) (holding that newspapers must comply with labor laws); Cohen v. Cowles Media Co., 501 U.S. 663, 672 (1991) (discussing a newspaper's contractual obligations).

111. *Cohen*, 501 U.S. at 669.

112. *See, e.g.*, Minneapolis Star Trib. Co. v. Comm'r of Revenue, 460 U.S. 575, 591–93 (1983) (holding that a use tax imposed on publications that used more than $100,000 per year in ink and paper a violation of the First Amendment).

113. *See, e.g.*, Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 UC DAVIS L. REV. 1183, 1186 (2016).

114. *See, e.g.*, Lina M. Khan & David E. Pozen, *A Skeptical View of Information Fiduciaries*, 133 HARV. L. REV. 497, 503–04 (2019) (observing that "[a] fiduciary with sharply opposed loyalties teeters on the edge of contradiction").

whether the entity satisfies the existing criteria but also whether those criteria can be adjusted to make the entity a fiduciary. If conducted within appropriate bounds, the process applies the general common law method to the question presented. Or, to modify the quoted statement from the Supreme Court, courts can apply "generally applicable" legal methods to the media without offending the First Amendment.

At some point, of course, simply slapping the label "fiduciary" on platforms does single them out for special treatment. The common law method allows for incremental development and changes in existing doctrine.[115] As Justice Holmes put it, common law judges "are confined from molar to molecular motions"[116]—small adjustments rather than large changes. And the proposition that the First Amendment does not preclude the application of general laws to the media but does limit the application of laws that single out the media for special treatment converts that observation into a statement about constitutional doctrine. We cannot assess whether the proposed application of antitrust, fiduciary, or common carrier laws to platforms is molecular rather than molar without examining the details, few of which proponents of these doctrinal developments spell out.[117]

C. Platform Liability for Utterances Not Protected by Free Expression

Most of the advocacy for adjusting the rules in place because of NIT focuses on the increased harm NIT enable. Platform liability for disseminating utterances that violate the rules in place could be justified on two grounds. Platforms enable speakers to inflict harm, becoming at least metaphorically accomplices in the harmful act. They sometimes enable people to inflict harm anonymously, making it difficult, if not impossible, to impose liability on the speaker.[118] Platform liability gives platforms an incentive to self-police before distributing constitutionally unprotected material.

1. Accomplice Liability

Traditional accomplice liability, outside the domain of free expression domain, arises when several conditions are met. What the accomplice does must

---

115. *See* Edward H. Levi, *An Introduction to Legal Reasoning*, 15 U. CHI. L. REV. 501,503 (1948) (describing legal reasoning as a process in which "the classification changes as the classification is made").

116. S. Pac. Co. v. Jensen, 244 U.S. 205, 221 (1917) (Holmes, J., dissenting).

117. Although I have quoted Holmes on the limits of common law judging, the constitutionalized principle would apply to statutes that adjusted fiduciary or antitrust law in media-specific ways. Statutes could make incremental changes in fiduciary law with effects on the media, but large changes with the effect or for the purpose of bringing platforms within the fiduciary category would have to be justified by applying either an intermediate or a strict standard of review.

118. Even if readily identified, the person actually making the unprotected utterance might be judgment proof where the preferred remedy is damages, as is typical in libel cases. This is not a problem distinctive to harm inflicted via NIT though, except insofar as NIT bring out from their hiding places a larger number of judgment proof offenders.

aid or encourage the unlawful act.[119] And more importantly, the accomplice must have a specified mental state—usually knowledge, but sometimes less than that—associated with her action. The usual example is the hardware store owner who sells rat poison to someone who uses it to kill a romantic rival. Simply doing so does not make the store owner an accomplice to murder. The picture changes, though, if the buyer says, "I'm thinking about killing someone. What's a good poison to use?" Here the store owner's knowledge might make her an accomplice. In another example, a wild-eyed person rushes into the store and exclaims, "sell me a big knife right away! I really need it!" The customer then uses the knife to kill someone standing outside the store. The owner might be an accomplice if she was reckless in selling the knife under the circumstances. As we will see, different mental elements for liability play a crucial role when we impose accomplice liability for statements unprotected by substantive free expression doctrine.

We know that traditional media can be held liable for harms caused by an expression that originates elsewhere. *New York Times v. Sullivan* was a libel case about a paid advertisement, but it was not at issue that the ad's authors, not a Times employee, wrote the allegedly libelous words.[120] Nor was the theory of liability predicated on some idea that by publishing the ad, the newspaper had "adopted" the words as its own.[121] The general rule, subject to variation around the edges, is that a person is liable for republishing someone else's libelous statements if the defamed person proves that the republication satisfies the applicable standards for liability, unless a specific privilege, such as that for fair reporting of statements made in official proceedings, attaches.[122]

The same can be said about true threats. Consider a radio talk-show host who recites over the air a true threat sent him by a listener. The threatening listener can be held liable for making a true threat, and it is likely that the talk-show host could be held liable as an accomplice without violating his right to free expression if the host knew that he was transmitting a true threat and perhaps even if the host was merely reckless in doing so.

Against the background that traditional media can be held liable as accomplices when they disseminate at least some types of unprotected speech, when the dissemination is done with the requisite mental state, does dissemination via NIT call for a modification of the existing elements of accomplice liability? On the face of things, it is not clear that those elements require modification. Professor Wu's example of accomplice liability—a president who directs (or perhaps only encourages) his followers to assault a critic—seems to me to satisfy the mental element requirement for accomplice

---

119. For an overview of the requirements for accomplice liability, *see* Sherif Girgis, *The Mens Rea of Accomplice Liability: Supporting Intentions*, 123 YALE L.J. 460, 465-68 (2013).

120. N.Y. Times Co. v. Sullivan, 376 U.S. 254, 256, 265-66 (1964).

121. *Id.* at 265-66.

122. *See* RESTATEMENT (SECOND) OF TORTS § 578 cmt. B & f (AM. L. INST. 1977).

liability.[123] At the least, there need to be more details about how the proposed accomplice liability would work in specific cases before we can think clearly about whether the current law of accomplice liability needs to be adjusted to cover actions by platforms that trouble the platforms' critics.[124]

Perhaps, though, accomplice liability pursuant to the rules applicable to traditional media would deter NIT platforms from disseminating material that the traditional media would disseminate in the face of those rules.

2. Incentive Effects

How should we think about the incentive effects of platform liability? Here we have models at hand in regulations adopted in Europe. These regulations apply to an expression such as true threats and to hate speech that is, in those nations, constitutionally unprotected.[125] Details vary, but the general contours are these: within a relatively short period after being notified that it has distributed some forms of unprotected speech—such as hate speech, a libelous statement, or the Holocaust lie—the platform must take it down by refraining from distributing it further and excising it from its searchable archives.[126] Failing to do so incurs monetary sanctions.[127]

Platform liability might lead platforms to create systems for identifying material that *might* trigger the take down obligation and refuse to disseminate it in the first place.[128] Such systems might use algorithms to identify material that might generate trouble for the platform and prevent that material from going onto the platform. And sometimes, the algorithms, or any other system, will lead the platform to refuse to disseminate material that is in fact constitutionally protected, such as true factual statements that injure reputation, opinions that

123. *See* Wu, *supra* note 1, at 17.

124. Imposing accomplice liability without requiring *any* mental state—knowledge, specific or general intent, recklessness—might well violate principles of free expression. I do not know whether many advocates for accomplice liability for platforms propose that form of liability.

125. American readers must be clear here not to confuse criticism of these legal systems for failing to provide protection to hate speech with criticism of the incentive effects of platform liability. It is not a criticism of platform liability that it gives incentives to refuse to distribute hate speech where such speech is not constitutionally protected.

126. *See, e.g.*, Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken [Network Enforcement Act] [BGB] [Civil Code], § 3, https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html (Ger.). The Act's Section 1 ("Scope") identifies "unlawful content" by reference to a list of statutory offenses that includes threats (Criminal Code § 126) and hate speech (Criminal Code § 166 ("revilement of religious faiths").

127. *See id.* at §§ 3, 4. The French Constitutional Council invalidated a similar law on proportionality grounds, emphasizing that its provisions, including an extremely short time period after notification for taking down proscribed material, gave platforms strong incentives to remove questionable content. Conseil constitutional [CC] [Constitutional Court] decision No. 2020-801DC, Jun. 18, 2020, J.O. 156 (Fr.).

128. A platform's decision to create such a system will depend upon the size of the monetary liability incurred if it lacks a system (or when material triggering liability is disseminated) and the cost of creating and operating the system.

injure reputation, or critical commentary on some factual assertions. This is, as before, the chilling effect of legal liability.

Suppose that a person can be held liable for distributing unprotected material, knowing it to be unprotected.[129] Even that rule will deter the distribution of some constitutionally protected material because the decision-maker tasked with imposing liability might mistakenly conclude that the speaker knew that the material was unprotected when the speaker, in actuality, did not know that. And generally, different mental elements will deter the distribution of different amounts of protected expression.

The rules in place are designed with their chilling effect in mind. That is, the amount of chilling associated with the rules in place is socially acceptable. The chilling effect of an actual knowledge standard might be socially acceptable, depending on circumstances, as might the chilling effect of a recklessness standard or a negligence one. Is there reason to think that the amount of chilling associated with platform liability for disseminating unprotected material is greater than that associated with doing so via the traditional media? Or, more precisely, is it enough greater than the latter to lead us to reconsider how much chilling is socially acceptable?[130]

Design details might matter, at least a bit. For example, does the platform face liability if it fails to take down materials when notified by an ordinary citizen, or only when it is notified by some recognized institution, whether governmental or otherwise? Identifying specific agencies whose notifications trigger the take-down obligation might reduce the disincentive effects. Narrowing the class of notifications that trigger the take-down requirement might do the same. And, to the extent that the agencies providing the notification have or develop some expertise in identifying unprotected material, the platform's uncertainty about whether the material is indeed unprotected is reduced. And, of course, the length of time before liability is triggered matters.[131]

I offer no answers to questions about the chilling effect of platform liability, no matter how designed. My point is that simply asserting that such liability will have a chilling effect is not a complete argument against platform liability

---

129. I find it difficult to imagine a sense in which the material is unprotected if liability could not attach in the posited example.

130. One set of answers to this question focuses on the rules in place. Suppose platform liability does increase the chilling effect of enforcing a specific rule in place. We might return the chilling effect to its prior level by changing the substantive rule by expanding the domain of protected expression. As before, such substantive responses have to be done on a category-by-category basis. Adjusted to take account of the proposition that here the response involves expanding rather than contracting the domain of protected expression, the arguments developed in Section III carry over here. *See supra* Section III.

131. Milder versions of liability for failing to take down material after notification can be devised. For example, liability might be imposed for failing to have in place an effective mechanism for taking unprotected material down after notification. The analytic issues associated with chilling effect arguments are the same though.

because the rules in place for traditional media have a chilling effect as well. Magnitudes matter.

And, finally, at this point, some basic concerns about constitutional law—not distinctively about free expression law—kick in. How much deference should a reviewing court give to a legislature's judgment that its enactment is constitutional? More particularly in the present context, how much deference should the court give to the legislature's judgment that what it has done satisfies a constitutional requirement of social acceptability?

Suppose a legislature enacts a statute imposing platform liability after, let us assume, a full and open discussion of its merits, including the size of the chilling effect. The enactment itself embodies a judgment by one important social institution that the size of the chilling effect is socially acceptable. Is it appropriate for a court assessing the statute's constitutionality to conclude that the size of the chilling effect is not socially acceptable? Or, again, more precisely, what standard of review should such a court apply to the question: is the chilling effect the legislature has found acceptable actually acceptable? That is how the most fundamental question about constitutional review arises in this context, and I have nothing new to add to its discussion.

CONCLUSION

This Essay has explored several lines of argument made in connection with the claim that the rise of NIT justifies adjusting the existing free expression doctrine. Its focus has been methodological, seeking to identify precisely how NIT might affect free expression doctrine without attempting to resolve normative questions, such as the relevance of material harm to threat liability, or empirical ones, such as the incremental harm associated with expression transmitted via NIT that increases the risk of law-breaking.

Yet, if the tenor of most discussions of these questions is that existing doctrine should be adjusted, the tenor of this Essay is to the contrary. I believe that the existing free expression doctrine is by and large adequate to deal with the challenges posed by NIT, especially when we acknowledge that doctrine is never frozen but allows for incremental change. And, I believe, incremental changes consistent with existing doctrine—tweaks—are generally enough to deal with these new challenges.[132] The First Amendment, in short, is not obsolete.

132. *See also* Massaro & Norton, *supra* note 1, at 1635 (favoring tweaking rather than toppling existing doctrine).