

**NLWJC - KAGAN**

**STAFF & OFFICE - D.C. CIRCUIT**

**BOX 008 - FOLDER 008 DC**

**Kagan - Forms [6]**

# FOIA MARKER

**This is not a textual record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.**

---

**Collection/Record Group:** Clinton Presidential Records

**Subgroup/Office of Origin:** Counsels Office

**Series/Staff Member:** Sarah Wilson

**Subseries:**

---

**OA/ID Number:** 14688

**FolderID:**

---

**Folder Title:**

Kagan - Forms [6]

**Stack:**

**V**

**Row:**

**13**

**Section:**

**2**

**Shelf:**

**11**

**Position:**

**3**

# The University of Chicago Law Review



---

Private Speech, Public Purpose:  
The Role of Governmental Motive in  
First Amendment Doctrine

*Elena Kagan*

© 1996 by The University of Chicago

---

Volume 63

Spring 1996

Number 2

CLINTON LIBRARY PHOTOCOPY

# Private Speech, Public Purpose: The Role of Governmental Motive in First Amendment Doctrine

Elena Kagan†

## INTRODUCTION

In one of the most frequently quoted passages of one of the most frequently cited First Amendment decisions, the Supreme Court declared that “the purpose of Congress . . . is not a basis for declaring [ ] legislation unconstitutional.”<sup>1</sup> Noting several hazards of attempting to ascertain legislative motive, the Court in *United States v O'Brien*<sup>2</sup> eschewed this endeavor in First Amendment cases, as well as in other constitutional adjudication. It was no task of the judiciary to discover or condemn “illicit legislative motive” relating to the freedom of speech; the question for courts was only whether a challenged statute, by its terms or in its application, had an “unconstitutional effect” on First Amendment freedoms.<sup>3</sup>

In keeping with this approach, most descriptive analyses of First Amendment law, as well as most normative discussions of the doctrine, have considered the permissibility of governmental regulation of speech by focusing on the effects of a given regulation. This focus on effects comes in two standard varieties. In one, the critical inquiry relates to the effect of a regulation on the speaker’s ability to communicate a desired message. In the other,

---

† Professor of Law, The University of Chicago (on leave 1995-96 to serve as Associate Counsel to the President). I am grateful to Mary Becker, David Currie, Richard Epstein, Richard Fallon, Stephen Gilles, Dan Kahan, Larry Lessig, Michael McConnell, Mark Ramseyer, Anne-Marie Slaughter, Geoffrey Stone, David Strauss, Cass Sunstein, and Laurence Tribe for extremely helpful comments. The Class of 1964 Fund and the Russell J. Parsons Faculty Research Fund provided financial support.

<sup>1</sup> *United States v O'Brien*, 391 US 367, 383 (1968).

<sup>2</sup> 391 US 367 (1968).

<sup>3</sup> *Id.* at 383, 385. In the *O'Brien* inquiry, the nature of the governmental interest asserted played an important role. See *id.* at 380-82; text accompanying notes 237-38. But the *O'Brien* Court cared not at all—or at least professed to care not at all—whether the asserted governmental interest matched, or even resembled, the actual interest underlying the enactment or enforcement of the legislation. *O'Brien*, 391 US at 383-85.

the critical inquiry relates to the effect of a regulation on the listener's ability to obtain information. In either case, however, what matters is the consequence of the regulation.

This Article shifts the focus from consequences to sources; I argue, notwithstanding the Court's protestations in *O'Brien*, that First Amendment law, as developed by the Supreme Court over the past several decades, has as its primary, though unstated, object the discovery of improper governmental motives. The doctrine comprises a series of tools to flush out illicit motives and to invalidate actions infected with them. Or, to put the point another way, the application of First Amendment law is best understood and most readily explained as a kind of motive-hunting.

This claim stands in need of much explanation, for as *O'Brien* indicates, even the attentive observer rarely catches a glimpse of the Court inquiring directly into governmental purpose. But assume for a moment that discovering impermissible motive stood as the Court's overriding object in the consideration of cases involving the First Amendment. Would the Court then charge itself with the task of dissecting and analyzing in each case the reasons animating the action of a governmental official or body? Not likely, for obvious reasons relating to the ease of legislatures' offering pretextual motives and the difficulty of courts' discovering the real ones. Would, then, the Court have to surrender its concern with motive? Not at all. The Court could construct and use objective tests to serve as proxies for a direct inquiry into motive. It could develop rules that operated, like certain burden-shifting mechanisms or presumptions, to counter the difficulties involved in determining motive and to enable the judiciary to make, if only indirectly, that determination.

The most important components of First Amendment doctrine—indeed, the very structure of that doctrine—serve precisely this function. If a court were to attempt to devise easily manageable rules for ferreting out impermissible governmental motives in the First Amendment context, it first would create a distinction between speech regulations that are content neutral and those that are content based. It then would develop a series of exceptions to that distinction in order to handle unusually suspicious kinds of content-neutral regulations and unusually trustworthy kinds of content-based restrictions. (This effort might give rise, for example, to the doctrine of so-called low-value speech.) It would add a division of great import between generally applicable regulations, only incidentally affecting speech, and regulations specifically targeted at expressive activity. If, in other words, a

court wished to construct a set of rules to determine impermissible motive in the First Amendment area, it might well devise the complex set of distinctions and categories currently governing First Amendment law. And conversely, if a court could determine governmental purpose directly, these rules, principles, and categories might all be unnecessary.

Courts, of course, rarely construct law in so deliberate a fashion; at least, the current Supreme Court—fractured, clerk-driven, and uninterested in theoretical issues as it is—rarely does so. The self-conscious rationalization and unification of bodies of law is not something to expect from the modern judiciary. So I do not mean to stake a claim that individual Justices, much less the Court as a whole, have set out intentionally to create a doctrinal structure that detects illicit motive by indirect means. The story I tell about purpose in the law does not depend on any assertion about the purpose of the Court. What I provide is simply a reading—I think the best reading—of the Court's First Amendment cases. I contend not that the Court self-consciously constructed First Amendment doctrine to ferret out improper motive, but that for whatever uncertain, complex, and unknowable reasons, the doctrine reads as if it had been so constructed.

I do not wish to overstate the case here, though perhaps I already have done so; I am not about to craft (yet another) all-encompassing—which almost necessarily means reductionist—theory of the First Amendment. First, what follows is primarily a descriptive theory; although I discuss its normative underpinnings, I make no claim that a sensible system of free speech should be concerned exclusively with governmental motivation. Second, even seen as descriptive, the theory has limits. Some aspects of First Amendment law resist explanation in terms of motive; other aspects, though explicable in terms of motive, are explicable as well by other means; and sometimes, the concern with governmental motive is itself intertwined with other apprehensions. And yet, all these qualifications notwithstanding, the concern with governmental motive remains a hugely important—indeed, the most important—explanatory factor in First Amendment law. If it does not account for the whole world of First Amendment doctrine, it accounts (and accounts alone) for a good part of it.

Section I of this Article introduces the discussion by using a recent case—*R.A.V. v City of St. Paul*<sup>4</sup>—to explore how a concern

---

<sup>4</sup> 505 US 377 (1992).

with impermissible motive underlies First Amendment doctrine. Section II moves backward to address the prior questions of what motives count as impermissible under the First Amendment and how such motives differ from legitimate reasons for restricting expression. Section III demonstrates how a wide range of First Amendment rules—indeed, the essential structure of the doctrine—are best and most easily understood as devices to detect the presence of illicit motive. Section IV concludes with some thoughts on the normative underpinnings of the Court's unstated, perhaps unrecognized, but still real decision to treat the question of motive as the preeminent inquiry under the First Amendment.

### I. AN EXAMPLE: THE PUZZLE OF *R.A.V. v CITY OF ST. PAUL*

Consider first the recent, important, and hotly debated Supreme Court decision of *R.A.V. v City of St. Paul*. The decision, invalidating a so-called hate speech ordinance, raises many questions about what counts, or should count, as the core concern of the First Amendment. An exploration of some of these questions shows in dramatic form the importance of governmental motive in the Court's First Amendment analysis.

*R.A.V.* arose from St. Paul's decision to charge a juvenile under its Bias-Motivated Crime Ordinance for allegedly burning a cross on the property of an African-American family. The ordinance declared it a misdemeanor for any person to "place[ ] on public or private property a symbol . . . which one knows or has reasonable grounds to know arouses anger, alarm or resentment in others on the basis of race, color, creed, religion or gender . . ." <sup>5</sup> In an effort to avoid constitutional problems, the Minnesota Supreme Court interpreted this statute narrowly to apply only to "fighting words" based on race, color, and so forth. <sup>6</sup> Courts long have considered fighting words to be unprotected expression—so valueless and so harmful that the government may ban them entirely without abridging the First Amendment. <sup>7</sup> The question thus raised by the state court's decision was whether St. Paul constitutionally could prohibit some, but not all, un-

<sup>5</sup> St. Paul, Minn, Legis Code § 292.02 (1990).

<sup>6</sup> See *In re R.A.V.*, 464 NW2d 507, 510-11 (Minn 1991). In *Chaplinsky v New Hampshire*, the United States Supreme Court defined "fighting words" as words "which by their very utterance inflict injury or tend to incite an immediate breach of the peace." 315 US 568, 572 (1942).

<sup>7</sup> See *Chaplinsky*, 315 US at 571-72.

protected speech—more specifically, fighting words based on race and the other listed categories, but no others.<sup>8</sup>

A majority of the Court, speaking through Justice Scalia, held that St. Paul could not take this action because it violated the principle of content neutrality. No matter that a city may ban all fighting words; it may not (as, the majority held, St. Paul did) ban only fighting words that address a particular subject or express a particular viewpoint. Although the category of fighting words is unprotected—although it has, “in and of itself, [no] claim upon the First Amendment”—the government does not have free rein to regulate selectively within the category.<sup>9</sup> Even wholly proscribable categories of speech are not “entirely invisible to the Constitution, so that they may be made the vehicles for content discrimination.”<sup>10</sup> To sustain such discrimination within categories of speech, just because the categories as a whole are proscribable, would be to adopt “a simplistic, all-or-nothing-at-all approach to First Amendment protection.”<sup>11</sup>

Three explanations for the Court’s decision offer themselves, the first two relating to different effects of the St. Paul ordinance, the last relating to its purpose. First, the Court might have held as it did because the St. Paul ordinance too greatly interfered with the opportunity of speakers to communicate their desired messages. Second, the Court might have reached its decision because the ordinance harmed the ability of the public—that is, the audience—to become exposed to a desirable range and balance of opinion. Third, the Court might have invalidated the ordinance because regardless how (or whether) it affected either speaker or audience, it stemmed from an improper purpose on the part of the government. Which of these three possibilities best explains the *R.A.V.* holding?

Not the first—not, that is, a perspective focusing on the speaker’s opportunity to engage in expression. As all of the Justices agreed, St. Paul could have enacted a statute banning all fighting words—a statute, in other words, imposing a more expansive restriction on speech than did the ordinance in ques-

---

<sup>8</sup> Four Justices believed there was no need to reach this question because the Minnesota Supreme Court had failed in its effort to limit the ordinance to fighting words, and the ordinance thus remained overbroad. See *R.A.V.*, 505 US at 413-14 (White concurring). The majority of the Court, however, declined to consider this argument, *R.A.V.*, 505 US at 381, and the dispute in the case focused on the question set out in the text.

<sup>9</sup> *R.A.V.*, 505 US at 386.

<sup>10</sup> *Id.* at 383-84.

<sup>11</sup> *Id.* at 384.

tion.<sup>12</sup> If St. Paul could have passed this broad limitation, silencing both the speakers affected by the actual ordinance and a great many others, then the flaw in the ordinance must have arisen from something other than its simple curtailment of expression. Consider here the views of Justice White, who wrote that "[i]t is inconsistent to hold that the government may proscribe an entire category of speech . . . but that the government may not [proscribe] a subset of that category . . ."<sup>13</sup> If expressive opportunities were the only constitutional interest, Justice White would be correct that the greater restriction includes the lesser. If he erred—if the greater does not, or does not always, include the lesser—it must be because of another interest.

Perhaps, then, the interest protected in *R.A.V.* is the interest of listeners in a balanced debate on public issues. The argument, initially plausible, goes as follows. The St. Paul ordinance on its face restricted speech on the basis of subject matter; fighting words based on race, but not on other topics, fell within its coverage. More, and more nefariously, the ordinance discriminated in its operation on the basis of viewpoint; the law effectively barred only the fighting words that racists (and not that opponents of racism) would wish to use.<sup>14</sup> The ordinance, while not restricting a great deal of speech, thus restricted speech in a way that skewed public debate on an issue by limiting the expressive opportunities of one side only. The reason the St. Paul ordinance posed a greater constitutional difficulty than a ban on all fighting words related to this skewing effect; the ordinance ensured that listeners would confront a distorted debate, thus interfering with "the thinking process of the community."<sup>15</sup>

---

<sup>12</sup> Id at 383-84; id at 401 (White concurring); id at 415-16 (Blackmun concurring); id at 417-18 (Stevens concurring).

<sup>13</sup> Id at 401 (White concurring).

<sup>14</sup> The question whether the St. Paul ordinance, in operation, discriminated on the basis of viewpoint divided the Justices, as it has divided commentators. Contrast the majority opinion, *R.A.V.*, 505 US at 391-92 (holding that ordinance was viewpoint discriminatory), with the concurring opinion of Justice Stevens, id at 434-35 (arguing that ordinance was not viewpoint discriminatory). Contrast also Cass R. Sunstein, *On Analogical Reasoning*, 106 Harv L Rev 741, 762-63 & n 78 (1993) (no viewpoint discrimination), with Elena Kagan, *The Changing Faces of First Amendment Neutrality: R.A.V. v St. Paul, Rust v Sullivan, and the Problem of Content-Based Underinclusion*, 1992 S Ct Rev 29, 69-71 (viewpoint discrimination). It is not necessary to resolve this issue here. If the ordinance, in application, did not amount to viewpoint discrimination, then the rationale based on skewing effects becomes much weaker. To best present the claim that the effects of the ordinance justify the *R.A.V.* decision, I assume the ordinance was viewpoint discriminatory.

<sup>15</sup> Alexander Meiklejohn, *Political Freedom: The Constitutional Powers of the People*

But on close inspection, this argument appears contrived. Even if the St. Paul statute distinguished between viewpoints on racial issues, several considerations, detailed here in ascending order of importance, suggest that the statute would not dangerously have distorted public debate.

First, the Court repeatedly has suggested that the expressive content of fighting words is, in Justice White's words, "worthless" or "*de minimis*."<sup>16</sup> If this understanding is accepted, a concern with the distortion of public debate in a case like *R.A.V.* looks awkward, even wholly misplaced. Assuming that fighting words have no expressive value—that they are not a part of public debate because not in form or function true communication—then the restriction of some fighting words, even if all on one side, cannot easily be thought to distort discussion. True, a law of this kind subtracts from one side only, but it subtracts a thing valued at zero and thus cannot change the essential equation.

I do not mean to claim that the distortion argument has no meaning in a sphere of unprotected expression. For one thing, any restriction on racist (but only racist) fighting words inevitably will chill racist (but only racist) speech outside the fighting words category; the chilling effect of such a regulation thus will cause some distortion in the realm of protected expression. More important, though courts often claim that fighting words and other unprotected speech have no expressive content or function, these claims ought not to be taken at face value. As Justice Scalia noted, "sometimes [fighting words] are quite expressive indeed."<sup>17</sup> Claims to the contrary serve as shorthand for a complex calculation that the harms of such speech outweigh their contribution to the sphere of expression. Thus, the subtraction of fighting words from one side of a debate is the subtraction not of a void, but of something quite tangible. And yet, in such a case, the concern with skewing the deliberative process continues to ring oddly, as it might if a law prevented one side of a debate from throwing brickbats at the other (an activity that also might be "expressive indeed"). A law of this kind would be unconstitu-

---

27 (Harper 1960) (emphasis omitted). Geoffrey Stone made this argument with respect to content-based regulation generally in *Content Regulation and the First Amendment*, 25 *Wm & Mary L Rev* 189, 198 (1983). I used the argument with respect to *R.A.V.* in *Kagan*, 1992 S Ct Rev at 63-64 (cited in note 14). As will become clear, I now find the argument unpersuasive, both in its application to *R.A.V.* and more broadly.

<sup>16</sup> *R.A.V.*, 505 US at 400 (White concurring).

<sup>17</sup> *R.A.V.*, 505 US at 385.

tional, but there is something peculiar in saying that this is because the law harms the thinking process of the community.

Second, even putting aside this objection, the skewing effect in a case like *R.A.V.* is very modest. Racists can continue to communicate their message in many ways; we need have no worry that the St. Paul ordinance will excise the idea of racism from public discourse, or indeed that the ordinance will noticeably cut into the idea's incidence. Of course, a flat rule, excluding case-specific inquiry of this kind, may have benefits. Judges, some will say, cannot reliably determine whether a given viewpoint-based law works only a modest distortion: the matter is one of degree and difficult to measure; perhaps, for example, the particular means restricted, though apparently modest, constitute the most effective way of delivering the message.<sup>18</sup> But this insight, and the preference for rule-based approaches that goes with it, cannot explain the *R.A.V.* decision. Within a sphere of unprotected speech (such as fighting words or obscenity), the most accurate generalization is that viewpoint distinctions will not significantly distort public discourse. Were skewing effects all that mattered, the *R.A.V.* Court thus would have established a bounded exception to the usual rule against viewpoint discrimination, applying in spheres of unprotected expression.

Finally, the notion of a skewing effect, as an explanation of *R.A.V.* or any other case, rests on a set of problematic foundations. The argument assumes that "distortion" of the realm of ideas arises from—and only from—direct governmental restrictions on the content of speech. But distortion of public discourse might arise also (or instead) from the many rules of property and other law that, without focusing or intending to focus on any particular speech, determine who has access to expressive opportunities.<sup>19</sup> If there is an "overabundance" of an idea in the absence of direct governmental action—which there well might be when compared with some ideal state of public debate—then action disfavoring that idea might "unskew," rather than skew, public discourse.<sup>20</sup> Suppose, for example, that racists control a

<sup>18</sup> See Stone, 25 *Wm & Mary L Rev* at 225-27 (cited in note 15).

<sup>19</sup> See Cass R. Sunstein, *Democracy and the Problem of Free Speech* 178-79 (Free Press 1993); Owen M. Fiss, *Why the State?*, 100 *Harv L Rev* 781, 786-87 (1987).

<sup>20</sup> See Sunstein, *Democracy and the Problem of Free Speech* at 178-79 (cited in note 19); Fiss, 100 *Harv L Rev* at 786-87 (cited in note 19). The formulation in the text suggests that there is such a thing as an ideal speech environment and may suggest, further, that we can describe its appearance. For doubts as to whether it is possible to provide an account of an optimal speech market, see David A. Strauss, *Rights and the System of*

disproportionate share of the available means of communication; then, a law like St. Paul's might provide a corrective.

A court well might—as the *R.A.V.* Court did—refuse the government the power to provide this corrective, but to do so, the court must discard a rationale focused purely on effects and adopt a rationale focused on motive. In denying the government a power of this kind, a court effectively determines that the “appropriate” distribution of speech is the distribution existing prior to direct governmental action. This determination, as I have noted, cannot be based on the view that the “preregulatory” distribution represents some platonic ideal of public discourse. It must be based on the view that whatever the existing state of affairs, direct restrictions, such as the St. Paul law, probably would worsen matters. And this thinking—the use of a presumption that governmental regulation will exacerbate, rather than ameliorate, distortion—is most naturally viewed as arising from a concern with the motives that underlie the regulation.<sup>21</sup> The worry in a case like *R.A.V.* is not with skewing effects per se; the fear of skewing effects depends upon, and becomes meaningless without, the fear that impermissible considerations—call them for now “censorial” or “ideological” considerations—intruded on the decision to restrict expression.

The *R.A.V.* Court made this concern about illegitimate, censorial motives unusually evident in its opinion, all but proclaiming that sources, not consequences, forced the decision. The First Amendment, the majority stated, “prevents government from proscribing speech . . . because of disapproval of the ideas expressed.”<sup>22</sup> And again: “The government may not regulate [speech] based on hostility—or favoritism—towards the underlying message expressed.”<sup>23</sup> The Court maintained that the structure of the ordinance—the subject-matter distinctions apparent on its face, the viewpoint distinctions apparent in operation—suggested illicit motive: “the nature of the content discrimination,” in the Court’s view, posed a “realistic possibility that of-

---

*Freedom of Expression*, 1993 U Chi Legal F 197, 205-07. The argument here does not rest on an ability to define fully an ideal state of debate; it rests only on the claim that the distribution of speech prior to direct governmental regulation need not, and usually will not, constitute such an ideal, or even something close to it.

<sup>21</sup> The presumption also might be thought to arise from a view of governmental incapacity to promote a healthier or more balanced speech market. I reject this alternative explanation in Section III.A, arguing that even if this general incapacity existed (which I doubt), it would provide insufficient reason to adopt the presumption.

<sup>22</sup> *R.A.V.*, 505 US at 382.

<sup>23</sup> *Id.* at 386.

ficial suppression of ideas [was] afoot."<sup>24</sup> And going beyond the structure of the law, the Court found that "comments and concessions" made by St. Paul in the case "elevate[d] the possibility to a certainty" that St. Paul was "seeking to handicap the expression of particular ideas" because of hostility toward them.<sup>25</sup>

Indeed, half hidden beneath a swirl of doctrinal formulations, the crux of the dispute between the majority and the concurring opinions concerned the proper understanding of St. Paul's motive in enacting its hate-speech law.<sup>26</sup> The majority understood this motive as purely censorial—a simple desire to blot out ideas of which the government or a majority of its citizens disapproved. The concurring Justices saw something different: an effort by the government, divorced from mere hostility toward ideas, to counter a severe and objectively ascertainable harm caused by (one form of) an idea's expression.<sup>27</sup> In part, this different understanding of motive emerged from a different view of the structure of the ordinance: in arguing that the ordinance did not discriminate on the basis of viewpoint, Justice Stevens suggested that the Court need not fear illicit purpose.<sup>28</sup> In part, the divergent interpretations of St. Paul's purpose reflected varying levels of sensitivity to the harms such speech causes.<sup>29</sup> In any event, the dispute was clear. "[T]his case does not concern the official suppression of ideas," said Justice White, but only a reasonable response to "pressing public concerns."<sup>30</sup> And Justice Stevens agreed that the ordinance had its basis not in "censorship," but in "legitimate, reasonable, and neutral justifications."<sup>31</sup>

---

<sup>24</sup> Id at 390.

<sup>25</sup> Id at 394. The Court noted especially a statement in St. Paul's Supreme Court brief that the purpose of the law was to show that the prohibited speech "is not condoned by the majority." Id at 392-93 (citations omitted).

<sup>26</sup> For a similar point, see Akhil Reed Amar, *The Case of the Missing Amendments: R.A.V. v. City of St. Paul*, 106 Harv L Rev 124, 146-51 (1992).

<sup>27</sup> The line between regulation based on hostility and regulation based on harm may be exceedingly fine. I discuss this distinction further in Section II.C.

<sup>28</sup> See *R.A.V.*, 505 US at 435 (Stevens concurring).

<sup>29</sup> Compare *R.A.V.*, 505 US at 393 ("St. Paul has not singled out an especially offensive mode of expression—it has not, for example, selected for prohibition only those fighting words that communicate ideas in a threatening (as opposed to a merely obnoxious) manner."), with id at 408 (White concurring) ("A prohibition on fighting words . . . is a ban on a class of speech that conveys an overriding message of personal injury and imminent violence, . . . a message that is at its ugliest when directed against groups that have long been targets of discrimination.").

<sup>30</sup> *R.A.V.*, 505 US at 411, 407 (White concurring).

<sup>31</sup> Id at 434, 416 (Stevens concurring).

The *R.A.V.* decision thus serves as a stark example of the importance of governmental motive in the Court's First Amendment analysis. Here, a debate about motive occurred in the open, revealing how a desire to punish impermissible purpose may explain and animate the Court's elaboration of doctrine. In the usual case, no such discussion occurs, but still the motive inquiry retains its power. The concern with filtering out illicit motive, though in these cases hidden, determines the content of the categories and rules that constitute First Amendment doctrine.

## II. THE CONCEPT OF IMPERMISSIBLE MOTIVE

What exactly does it mean to say that an effort to filter out impermissible motive animates and explains First Amendment doctrine? In part, the question will have its answer only after I show how particular categories and rules of First Amendment law reflect a concern with governmental motive. But before taking on that task, I must discuss in a more general way the nature of the concern with motive and the kinds of legal tools suited to address it. In this Section, I first compare an approach to the First Amendment focused on motive with two approaches focused on effects. I next attempt to define and delimit what motives count as improper under the First Amendment. Here, I describe the concept of impermissible motive operative in the doctrine, while deferring to Section IV a discussion of why this concept might have become central. Finally, I examine methods a legal system can use to address the question of impermissible motive, given the difficulties of proof (and, some might say, the problems of coherence) such an inquiry raises. In much of this Section's discussion, the reader will hear echoes of *R.A.V.*, as the concerns and strategies of the Court in that case assume a more general shape and structure.

### A. Three Perspectives on the First Amendment

In recent scholarship, a trend has developed to distinguish between two approaches to the First Amendment, which are sometimes complementary but often conflicting.<sup>32</sup> According to

---

<sup>32</sup> See Strauss, 1993 U Chi Legal F at 199-202 (cited in note 20) (contrasting an approach based on speakers' rights to a structural approach); Fiss, 100 Harv L Rev at 785-86 (cited in note 19) (contrasting an autonomy principle to a public debate principle); Robert Post, *Meiklejohn's Mistake: Individual Autonomy and the Reform of Public Discourse*, 64 U Colo L Rev 1109, 1132-33, 1136-37 (1993) (contrasting an autonomy theory to a collectivist theory).

this scholarly scheme, one conception of the First Amendment focuses on expanding the expressive opportunities open to speakers, whereas another focuses on improving the sphere of discourse encountered by the public "audience." To these two conceptions, which turn on different effects of speech regulation, I here juxtapose a third, which turns on the regulation's reasons.

The first approach—call it the "speaker-based" model—understands the primary value of the First Amendment to reside in its conferral of expressive opportunities on would-be communicators. A system of free expression, in allowing individuals to communicate their views, enhances their "autonomy" or "self-respect" or "self-development" or other (equally amorphous but desirable) human quality.<sup>33</sup> Under this theory, any limitation of expressive opportunities constitutes a harm because it interferes with some speaker's ability to communicate to others and with the benefit that speaker thereby derives. Moreover, the greater the limitation on speech, the greater the harm; under this theory, a broad restriction always poses greater constitutional concerns (because it interferes with more expressive activity) than a narrow one. Quantity, in other words, is of the essence; as one proponent of this model has stated, First Amendment doctrine should concern itself with how much a law "reduces the sum total of information or opinion disseminated."<sup>34</sup>

By contrast, the second approach to the First Amendment—call it the "audience-based" model—focuses on the quality of the expressive arena. A system of free expression, under this theory, has value because it enables the public—the audience for the speech—to arrive at truth and make wise decisions, especially about matters of public import.<sup>35</sup> In order best to fulfill this function, a system of free expression should promote not speech alone, but speech of a certain kind and mixture. Rich public debate is the goal; the concern is the expressive realm as a whole, rather than each opportunity for expression. Under this theory, restrictions on speech pose more or less danger depending not on

---

<sup>33</sup> For versions of this approach, see C. Edwin Baker, *Human Liberty and Freedom of Speech* 47-69 (Oxford 1989); Martin H. Redish, *The Value of Free Speech*, 130 U Pa L Rev 591 (1982); David A.J. Richards, *Free Speech and Obscenity Law: Toward a Moral Theory of the First Amendment*, 123 U Pa L Rev 45, 59-70 (1974).

<sup>34</sup> Martin H. Redish, *The Content Distinction in First Amendment Analysis*, 34 Stan L Rev 113, 128 (1981).

<sup>35</sup> See Meiklejohn, *Political Freedom* at 19-28 (cited in note 15); Sunstein, *Democracy and the Problem of Free Speech* at 53-77 (cited in note 19); Fiss, 100 Harv L Rev at 787-94 (cited in note 19).

the sum total of speech prevented, but on the extent to which the restrictions distort or impoverish the realm of discourse. Further, some restrictions on speech are preferable to none, given that some enhance public discussion—for example, by preventing a few voices from drowning out others.<sup>36</sup> The purpose of the First Amendment—the goal against which regulation must be measured—is the protection of what Alexander Meiklejohn called the public's "thinking process" from injury or "mutilation."<sup>37</sup>

The differences between these two approaches are captured in another of Meiklejohn's sayings. "What is essential," he wrote in support of the audience-based model, "is not that everyone shall speak, but that everything worth saying shall be said."<sup>38</sup> Place to one side Meiklejohn's view of the essential, and what remains in this aphorism is the core divergence between the models: one focuses on the effects of regulation on who speaks, the other on the effects of regulation on what is spoken. But this statement of the models' disparity reveals also their likeness: both make critical an action's consequences.

The third approach to the First Amendment—call it the "government-based" or "motive-based" model—claims that what is essential is not the consequences of a regulation but the reasons that underlie it.<sup>39</sup> The point of attention is neither the

---

<sup>36</sup> It is possible to contend that direct speech regulation never will serve the goal of rich public debate, or that it will so rarely serve that goal as to allow courts to assume that it never will do so. In some sense, the classic "marketplace of ideas" theory takes this position: the goal is to have a realm of discourse that leads to truth, but the means is a laissez-faire system. For this view to make sense, however, there must be a reason to think that the absence of regulation will lead to better results—here, a truth-producing market—than the allowance of regulation designed (or purportedly designed) to achieve this object. The most powerful such reason has to do with the government's other motives for curtailing speech. (I reject in Section III.A an alternative reason, relating to the sheer incapacity of government to improve the speech market.) In this sense, any argument that advances the quality of debate as a goal, but assumes that an absence of regulation will best achieve this object, should be seen as an argument about governmental motive. The point here resembles the one made in considering *R.A.V.*: there is little reason to think a speech restriction necessarily will skew, rather than balance, public debate in the absence of a concern about governmental motive.

<sup>37</sup> Meiklejohn, *Political Freedom* at 27 (cited in note 15) (emphasis omitted).

<sup>38</sup> *Id.* at 26.

<sup>39</sup> Of the three approaches, the government-based approach is least represented in the literature. Frederick Schauer is the principal proponent of an approach of this kind; he has emphasized the danger of illicit governmental motive as part of a normative defense of providing heightened protection for expression. See Frederick Schauer, *Free speech: a philosophical enquiry* 80-86 (Cambridge 1982). Geoffrey Stone and Cass Sunstein both have included considerations of motive in broader analyses of First Amendment law. See Stone, 25 *Wm & Mary L Rev* at 227-33 (cited in note 15); Sunstein, *Democracy and the Problem of Free Speech* at 154-59 (cited in note 19). Finally, advocates of

the sum total of speech prevented, but on the extent to which the restrictions distort or impoverish the realm of discourse. Further, some restrictions on speech are preferable to none, given that some enhance public discussion—for example, by preventing a few voices from drowning out others.<sup>36</sup> The purpose of the First Amendment—the goal against which regulation must be measured—is the protection of what Alexander Meiklejohn called the public's "thinking process" from injury or "mutilation."<sup>37</sup>

The differences between these two approaches are captured in another of Meiklejohn's sayings. "What is essential," he wrote in support of the audience-based model, "is not that everyone shall speak, but that everything worth saying shall be said."<sup>38</sup> Place to one side Meiklejohn's view of the essential, and what remains in this aphorism is the core divergence between the models: one focuses on the effects of regulation on who speaks, the other on the effects of regulation on what is spoken. But this statement of the models' disparity reveals also their likeness: both make critical an action's consequences.

The third approach to the First Amendment—call it the "government-based" or "motive-based" model—claims that what is essential is not the consequences of a regulation but the reasons that underlie it.<sup>39</sup> The point of attention is neither the

---

<sup>36</sup> It is possible to contend that direct speech regulation never will serve the goal of rich public debate, or that it will so rarely serve that goal as to allow courts to assume that it never will do so. In some sense, the classic "marketplace of ideas" theory takes this position: the goal is to have a realm of discourse that leads to truth, but the means is a laissez-faire system. For this view to make sense, however, there must be a reason to think that the absence of regulation will lead to better results—here, a truth-producing market—than the allowance of regulation designed (or purportedly designed) to achieve this object. The most powerful such reason has to do with the government's other motives for curtailing speech. (I reject in Section III.A an alternative reason, relating to the sheer incapacity of government to improve the speech market.) In this sense, any argument that advances the quality of debate as a goal, but assumes that an absence of regulation will best achieve this object, should be seen as an argument about governmental motive. The point here resembles the one made in considering *R.A.V.*: there is little reason to think a speech restriction necessarily will skew, rather than balance, public debate in the absence of a concern about governmental motive.

<sup>37</sup> Meiklejohn, *Political Freedom* at 27 (cited in note 15) (emphasis omitted).

<sup>38</sup> *Id.* at 26.

<sup>39</sup> Of the three approaches, the government-based approach is least represented in the literature. Frederick Schauer is the principal proponent of an approach of this kind; he has emphasized the danger of illicit governmental motive as part of a normative defense of providing heightened protection for expression. See Frederick Schauer, *Free speech: a philosophical enquiry* 80-86 (Cambridge 1982). Geoffrey Stone and Cass Sunstein both have included considerations of motive in broader analyses of First Amendment law. See Stone, 25 *Wm & Mary L Rev* at 227-33 (cited in note 15); Sunstein, *Democracy and the Problem of Free Speech* at 154-59 (cited in note 19). Finally, advocates of

speaker nor the audience, but the governmental actor standing in the way of the communicative process. Under this model, an action may violate the First Amendment because its basis is illegitimate, regardless of the effects of the action on either the sum of expressive opportunities or the condition of public discourse. Conversely, an action may comport with the First Amendment because legitimate reasons underlie it, again regardless of its range of consequences. The critical inquiry concerns what lies behind, rather than what proceeds from, an exercise of governmental power.<sup>40</sup>

The divide between the model based on motive and the models based on effects can be overstated.<sup>41</sup> One reason for First Amendment law to worry about governmental motive is itself consequential in nature; it refers to the predictable tendency of improperly motivated actions to have certain untoward effects.<sup>42</sup> To say this is not to collapse the distinction I have offered. First, the government- or motive-based model may emerge as well from nonconsequential, deontological considerations, relating to the stance or attitude we expect the government to adopt in relation to its citizens. Second, the government-based approach—even if in an ultimate sense inspired by a concern for consequen-

---

so-called listener-autonomy theories of the First Amendment, such as David Strauss and (at one time) Thomas Scanlon, may be engaged in a form of motive analysis, in that they appear to contest the legitimacy not of speech regulations themselves, but of certain (autonomy-infringing) justifications for them. See David A. Strauss, *Persuasion, Autonomy, and Freedom of Expression*, 91 Colum L Rev 334, 353-60 (1991); Thomas Scanlon, *A Theory of Freedom of Expression*, 1 Phil & Pub Aff 204, 209 (1972). The authors of such theories, however, do not speak in terms of motive and might well contest my characterization. See Strauss, 1993 U Chi Legal F at 201 (cited in note 20) (asserting that listener autonomy theories fall under what I have called an "audience-based" approach).

<sup>40</sup> As should be obvious by now, I make no distinction between such terms as "purpose," "intent," "motive," "basis," and "reason." The Court has used these terms interchangeably, both in First Amendment jurisprudence and elsewhere; in *O'Brien*, for example, the Court treated the terms "motive" and "purpose" as synonymous. See 391 US at 383. See also David A. Strauss, *Discriminatory Intent and the Taming of Brown*, 56 U Chi L Rev 935, 951 (1989) (noting the interchangeable use of these terms in equal protection law). Moreover, attempts by scholars to distinguish among these terms have proved unhelpful. See John Hart Ely, *Legislative and Administrative Motivation in Constitutional Law*, 79 Yale L J 1205, 1217-21 (1970) (criticizing such efforts).

<sup>41</sup> So too can the divide between the two models based on effects. For example, protection of audience-based interests demands that the government accord some substantial rights to individual speakers. Conversely, protection of speaker-based interests may demand some attention to the condition of public discourse, to prevent opportunities for expression from becoming purely formal and ineffective. A set of complicated relations exists among all these models, as well as among the concerns that underlie them.

<sup>42</sup> For further discussion of the points made in this paragraph, see text accompanying notes 257-78.

ces—very often will lead to different doctrinal rules, producing different results, than an approach that focuses on effects, whether on the speaker or the audience. Still, this analytic relationship between a motive-based approach and effects-based approaches should not be disregarded; I do not want to suggest that these approaches exist hermetically sealed from each other.

Further, the motive-based model and the effects-based models can operate in confluence with each other, except in their starkest forms—each contributing something to First Amendment doctrine. A body of law predominantly concerned with effects (whether on the speaker or audience) can make some place for considerations of governmental purpose; so too, but conversely, for a body of law predominantly concerned with motive. I have no doubt that current doctrine responds, in some manner and at some times, to all the concerns I have mentioned. The government-based approach does not wholly exclude the others.

The delineation of the three approaches, however, remains important. The approaches often will point in divergent directions, prescribing both different rules of law and different outcomes. And the pattern of decisions where such conflicts take place says much about the concerns that drive the law of free speech. To prepare the way for showing that among the potential concerns, illicit motive takes pride of place, I turn now to the meaning of illicit motive in First Amendment analysis.

## B. Defining Impermissible Motive

Assuming for now that First Amendment law constitutes an attempt to flush out impermissible motives, what motives count as impermissible? The Court has not fully addressed, much less resolved, this question. Despite the proscription in *O'Brien*, the Court sometimes has probed the government's reasons for restricting expression;<sup>43</sup> too, the Court has articulated several statements of First Amendment principle that sound in terms of

---

<sup>43</sup> The Court most recently inquired into legislative motive in *R.A.V. and Turner Broadcasting System, Inc. v FCC*, 114 S Ct 2445, 2461-62 (1994). A much earlier example of such an examination appears in *Grosjean v American Press Co.*, 297 US 233, 250 (1936). In cases involving executive action, the Court routinely speaks in terms of motive. For example, in addressing a First Amendment challenge brought by a discharged employee of the government, the Court will ask whether the government fired the employee because it disapproved of her expression. See, for example, *Connick v Myers*, 461 US 138, 143-46 (1983). The Court apparently sees the examination of motive in such cases as different in kind from—and less problematic than—the examination of the motives underlying legislation.

motivation.<sup>44</sup> But the Court, as the edict in *O'Brien* shows, usually has hesitated to discuss the issue of illicit motive in any detail or with any directness. The effort to define the concept of illicit purpose operative in First Amendment law thus must depend on a fair amount of extrapolation, as well as on a "reading backwards" from the doctrines discussed later in this Article.

The definition of illicit motive that this effort reveals is in certain respects imprecise and conceptually puzzling. It is not necessary, given my purpose, to untangle all the complexities this definition raises; what matters for this Article is that the doctrine emerges from an understanding of illicit motive, however inexact or enigmatic. This Section, then, provides only a sketch of the definitional issues. I first lay out the concept of impermissible motive evident in the law, in part describing it in terms of a neutrality principle. I then consider both a limitation on and an objection to the definition I have offered, and I finally compare that definition with some alternatives.

#### 1. A definition.

Consider the following snapshot of impermissible motives for speech restrictions. First, the government may not restrict expressive activities because it disagrees with or disapproves of the ideas espoused by the speaker; it may not act on the basis of a view of what is a true (or false) belief or a right (or wrong) opinion.<sup>45</sup> Or, to say this in a slightly different way, the government cannot count as a harm, which it has a legitimate interest in preventing, that ideas it considers faulty or abhorrent enter the public dialogue and challenge the official understanding of acceptability or correctness. Second, though relatedly, the government may not restrict speech because the ideas espoused threaten officials' own self-interest—more particularly, their tenure in office.<sup>46</sup> The government, to use the same construction as above, cannot count as a harm, which it has a legitimate interest in preventing, that speech may promote the removal of

---

<sup>44</sup> See notes 45, 48-49.

<sup>45</sup> See *City Council v Taxpayers for Vincent*, 466 US 789, 804 (1984) (asking whether a law "was designed to suppress certain ideas that the City finds distasteful"); *Consolidated Edison Co. v Public Service Commission*, 447 US 530, 536 (1980), quoting *Niemotko v Maryland*, 340 US 268, 282 (1951) (Frankfurter concurring) (asking whether speech was barred "merely because public officials disapprove the speaker's views"); Stone, 25 Wm & Mary L Rev at 227-28 (cited in note 15).

<sup>46</sup> See Stone, 25 Wm & Mary L Rev at 228 (cited in note 15); Sunstein, *Democracy and the Problem of Free Speech* at 155 (cited in note 19).

incumbent officeholders through the political process. Third, and as a corollary to these proscriptions, the government may not privilege either ideas it favors or ideas advancing its self-interest—for example, by exempting certain ideas from a general prohibition.<sup>47</sup> Justice Scalia summarized these tenets in *R.A.V.*: “The government may not regulate [speech] based on hostility—or favoritism—towards the underlying message expressed.”<sup>48</sup>

To this statement of illicit motive, one further gloss must be added: the government may not limit speech because other citizens deem the ideas offered to be wrong or offensive<sup>49</sup>—or for that matter, because they see the ideas as threatening to incumbent officials. This ban echoes those just stated, except for the identity of the party (above the government, now the public) that disapproves the ideas; the theory is that this substitution of party name should make no constitutional difference. Some of course may argue that restrictions based on public dislike boast a democratic legitimacy separating them from restrictions based on governmental hostility. But this distinction falters on the difficulty of disentangling the actions of officials from the desires of constituents. When the government acts, its reasons for doing so usually reflect the views of some part of the public. Distinguishing between public and governmental hostility thus seems hopeless as a practical matter. Further, the distinction shatters on the Court’s longstanding view that the First Amendment protects

---

<sup>47</sup> See John Hart Ely, *Flag Desecration: A Case Study in the Roles of Categorization and Balancing in First Amendment Analysis*, 88 Harv L Rev 1482, 1506-07 (1975); Stone, 25 Wm & Mary L Rev at 227-28 (cited in note 15).

<sup>48</sup> 505 US at 386. See also *Young v American Mini Theatres, Inc.*, 427 US 50, 67 (1976) (“[R]egulation of communication may not be affected by sympathy or hostility for the point of view being expressed by the communicator.”).

<sup>49</sup> See *Texas v Johnson*, 491 US 397, 414 (1989) (“If there is a bedrock principle underlying the First Amendment, it is that the Government may not prohibit the expression of an idea simply because society finds the idea itself offensive or disagreeable.”); *Hustler Magazine, Inc. v Falwell*, 485 US 46, 55 (1988), quoting *FCC v Pacifica Foundation*, 438 US 726, 745-46 (1978) (“The fact that society may find speech offensive is not a sufficient reason for suppressing it. Indeed, if it is the speaker’s opinion that gives offense, that consequence is a reason for according it constitutional protection.”). See also Stone, 25 Wm & Mary L Rev at 214-16 (cited in note 15); Sunstein, *Democracy and the Problem of Free Speech* at 155-56 (cited in note 19). The notion of offense is, of course, a tricky one. There is a fine line between offense at the content of ideas, to which I refer, and offense at the means by which those ideas are expressed. There is also a fine line between mere offense and emotional injury, in that a certain kind and degree of the former (the “offense” felt, for example, by the concentration camp survivors in Skokie) may constitute what society recognizes, or would wish to recognize, as the latter. Finally there is a complex relationship between offense at ideas (or any other sort of hostility toward ideas) and the entire range of harms those ideas cause. See text accompanying notes 60-66.

no less against majority oppression than against runaway government.<sup>50</sup> In keeping with this general view, the Court's conception of illicit motive must apply not only to officials but also to the public acting through them. Just as in equal protection law the government may not discriminate among persons on the basis of majoritarian biases,<sup>51</sup> so too in First Amendment law the government may not so distinguish among messages. The key principle with respect to motive is that the government may not limit speech on grounds of mere disapproval, no matter whose or how widely shared.

This principle leaves untouched many reasons to restrict expression; in this Article, I call these reasons "harm-based" and contrast them to the "ideological" reasons I have just discussed. The distinction raises difficult issues, which I address below, but to understand first its essential nature, a return to *R.A.V.* may prove helpful. Consider some different explanations for the St. Paul ordinance. First, the city may have enacted the statute to express its own or its citizenry's hatred of the ideas of racism, sexism, and so forth.<sup>52</sup> Alternatively, the city could have enacted the statute to prevent harms that it thought the covered speech posed to the community. Perhaps the city feared that the speech would cause some persons to suffer psychic trauma or other emotional harm; or that the speech would spark bloody public riots, because of strong popular resistance to it;<sup>53</sup> or that the speech

---

<sup>50</sup> See cases cited in note 49. See also *Kingsley Pictures Corp v Regents*, 360 US 684, 689 (1959) (stating that the First Amendment's "guarantee is not confined to the expression of ideas that are conventional or shared by a majority"). This view has its ancestry in the Framers' fear of majority factions. As James Madison wrote: "In our Governments the real power lies in the majority of the Community, and the invasion of private rights is chiefly to be apprehended, not from acts of Government contrary to the sense of its constituents, but from acts in which the Government is the mere instrument of the major number of the constituents." Letter from James Madison to Thomas Jefferson (Oct 17, 1788), in Robert A. Rutland, et al, eds, 11 *The Papers of James Madison* 298 (Virginia 1977). But see Akhil Reed Amar, *The Bill of Rights as a Constitution*, 100 Yale L J 1131, 1147-52 (1991) (suggesting that the First Amendment protects against government self-dealing, rather than majority tyranny). Amar's narrow understanding of the Amendment suggests, among other things, that no decision made by a properly selected jury—in, for example, a defamation case—would violate free speech principles.

<sup>51</sup> See, for example, *City of Cleburne v Cleburne Living Center, Inc.*, 473 US 432, 448 (1985); *Palmore v Sidoti*, 466 US 429, 433 (1984).

<sup>52</sup> Justice Scalia believed the city intended the law to serve just this function and accordingly savaged the city's motives. Asserting that the city enacted the ordinance to convey the majority's disapproval of an idea, Scalia wrote that "[t]he point of the First Amendment is that majority preferences must be expressed in some fashion other than silencing speech . . ." *R.A.V.*, 505 US at 392.

<sup>53</sup> This reason for restricting speech in one sense depends on popular hostility to ideas, which I have deemed an illegitimate reason for speech regulation: were it not for

would persuade listeners to engage in acts of race-based violence. The concept of impermissible motive I have described applies to the first of these explanations, but to none of the others. The first violates the principle that the government may not restrict speech on the basis of its own or the majority's view of what ideas are right or wrong, praiseworthy or shameful. The others do not violate this principle because they relate not to the message as message, but to the consequence of its expression; they stem not from ideological hostility, but from a perception of material harm.<sup>54</sup> In short and critically, they relate to harms that the government has a legitimate interest in preventing and obviously could act to prevent if not caused by expression.

Whenever hostility toward ideas as such (or the other impermissible factors of sympathy or self-interest) has played some part in effecting a restriction on speech, the restriction is irretrievably tainted; what has entered into the action commands its invalidation.<sup>55</sup> In contrast, when such factors have played no role—when the government has restricted ideas only as and when they bear harmful consequences—the government's purposes support sustaining the action. The critical inquiry is whether the government would have imposed the restriction in the absence of impermissible factors, solely on the basis of a neutral and legitimate evaluation of harm. Or to put the question in another way, it is whether the government would have treated (or did treat) identically ideas with which it disagreed, ideas with which it agreed, and ideas to which it was indifferent, to the extent those ideas caused the same harms.<sup>56</sup> This inquiry tests

---

popular hostility, the government would have no fear of riot, and thus would have no reason to restrict the expression. Nonetheless, there is a difference between restricting speech because of public hostility alone and restricting speech because this hostility will lead to bloodshed. In the former case, the government acts only to advance the majority's version of truth; in the latter, the government acts to avert violence, even if it is the majority's desire to impose its will that makes this action necessary.

<sup>54</sup> The Court recently drew this kind of distinction in explicit terms, asking whether "the Government seeks not to advance a legitimate regulatory goal, but to suppress unpopular ideas or information or manipulate the public debate." *Turner Broadcasting System, Inc. v FCC*, 114 S Ct 2445, 2458 (1994).

<sup>55</sup> For a similar understanding of the consequence of finding an impermissible motive, see Stone, 25 Wm & Mary L Rev at 229-30 (cited in note 15). Note that it should make no difference whether the impermissible motive has played a role on a conscious or unconscious level. If, as I argue, the injury is differential treatment based on prohibited considerations, the injury is not affected by the level of consciousness at which the considerations operated. See Strauss, 56 U Chi L Rev at 960 (cited in note 40).

<sup>56</sup> This test resembles the test proposed by many commentators to determine discriminatory intent in the equal protection context. See, for example, Paul Brest, *The Supreme Court, 1975 Term—Foreword: In Defense of the Antidiscrimination Principle*, 90 Harv L

whether the government regulated, even in part, on the basis of ideas as ideas, rather than on the basis of material harms.

This test is a measure of the neutrality or impartiality that the First Amendment often is said to command.<sup>57</sup> The First Amendment allows distinctions among speech on many bases. What it does not allow is classifications built on hostility or sympathy to ideas. The neutrality principle thus mirrors the doctrine of impermissible purpose. The government may classify speech to achieve legitimate governmental objects, such as the prevention of illegality or violence. But the government may not rest a classification, even in part, on the ground that some messages are worthier than others. Differences of this kind with respect to ideas must count as legally irrelevant. To say that "[t]here is an 'equality of status in the field of ideas,'"<sup>58</sup> is to say that the government cannot regulate speech for such impermissible reasons.

## 2. A limitation, an objection, and a comparison.

One question about the principle just articulated relates to its scope of operation: does the principle apply only when the government acts in its traditional role as regulator of private speech, or does it also apply (in either pure or diluted form) when the government performs the increasingly important functions of speaker, employer, and educator? In this Article, I discuss the issue of governmental motive only in relation to restrictions on private speech; except for a few words, I leave for another time the question how the understanding of improper motive I have described translates (or does not) into contexts in which the government itself performs speech functions. My thumbnail view is that the principle has greater relevance in these contexts than

---

Rev 1, 6-8 (1976); Ely, 79 Yale L J at 1266-68 (cited in note 40). David Strauss calls this the "reversing the groups" test because it asks "if the government would have made the same decision even if the races of those affected had been reversed." Strauss, 56 U Chi L Rev at 957 (cited in note 40). Similarly, the test in the First Amendment context might be called the "reversing the ideas" test, because it asks whether the government would have made the same decision if different ideas were affected. Strauss criticizes the test in equal protection law on the ground that it forces courts to make speculative, counterfactual determinations. To the extent the test asks courts to consider the question of "reversal" directly, this criticism is valid; my argument, offered in Section III, is that First Amendment doctrine relieves courts of this impossible task by providing rules that capture the gist of this inquiry in a concrete and easily administrable fashion.

<sup>57</sup> See, for example, *Carey v Brown*, 447 US 455, 462-63 (1980); *Police Department of Chicago v Mosley*, 408 US 92, 96 (1972).

<sup>58</sup> *Mosley*, 408 US at 96, quoting Meiklejohn, *Political Freedom* at 27 (cited in note 15).

might be thought, though less than when the government restricts private speech;<sup>59</sup> I also believe that the concept of illicit purpose should apply in these contexts even more strongly than it does, thus narrowing (though not eliminating) the importance for First Amendment analysis of the particular role the government is playing. But because I cannot defend these views in this Article, the key point here is one of limitation: the concept of impermissible motive I have described refers to the government in its capacity as regulator of private expression.

A second and, for my purposes, more important question concerns the coherence of the distinction I have drawn between motives based on harm and motives based on ideology—a distinction that might be viewed as possessing rhetorical appeal, but collapsing on deep reflection.<sup>60</sup> What is it, after all, to hate a message if not, and other than, to think the message causes injury? Perhaps opposition to speech on what I have termed “ideological” grounds—sheer hostility toward a message—does not exist as a real-world phenomenon. Perhaps such opposition always stems from, and thus reduces to, a conviction that the idea causes harms that the government has a legitimate interest in preventing.<sup>61</sup> If this is so, the distinction I have drawn might be said to rest only on the level of generality chosen to frame the critical question. Query 1: Why did officials restrict the expression? Answer 1: Because they disliked its message. Query 2: Why did the officials dislike its message? Answer 2: Because they believed the message caused material harm of a serious nature. If the distinction I have drawn depends on failing to ask the second question, then it seems a foundation too weak to support First Amendment doctrine.

This challenge is strong and the issue complex, but some examples indicate that the two kinds of motives, though closely interwoven, retain distinct characters. Assume that racist speech—or, to see the point from another perspective, assume that flag-burning—poses dangers: such speech may spark a riot,

---

<sup>59</sup> The law in this area is largely a mess, resisting any coherent understanding. If motive-based theory does not wholly explain the doctrine, neither does any other. See Kagan, 1992 S Ct Rev at 40-45 (cited in note 14).

<sup>60</sup> I assume that the distinction between motives based on harm and motives based on self-interest is not so mysterious. It seems clear that self-interest can counsel a speech restriction that an evaluation of harm (at least of harm that can be counted as harm) would not.

<sup>61</sup> See Elena Kagan, *Regulation of Hate Speech and Pornography After R.A.V.*, 60 U Chi L Rev 873, 880 (1993).

induce a violation of law, or cause emotional injury. Not everyone will measure or respond to these potential harms in the same manner. Persons will differ both in assessing the magnitude (indeed, the existence) of danger and in deciding what amount of danger will justify a restriction. And these divergent judgments about the harm the speech causes and the need to limit it rest in part on what I have said cannot count in the equation: the desire of persons, conscious or not, to suppress ideas that challenge (just because they challenge) and to privilege ideas that ratify (just because they ratify) their own belief systems.

So too we might explain other instances, past and present, of deciding when neutral interests counsel a restriction on speech. Consider the core cases of our free speech tradition, involving the questions whether speech opposing World War I or supporting communism threatened resistance to law or overthrow of the government.<sup>62</sup> Or consider the string of cases in the 1960s raising the issue whether civil rights protests would cause public riots.<sup>63</sup> Or consider this past year's debate as to whether right-wing talk radio provokes crimes of violence.<sup>64</sup>

As examples of this kind suggest, hostility toward speech (or its opposite) may affect the decision to regulate speech, separate from and independent of neutrally conceived harms.<sup>65</sup> Such hos-

---

<sup>62</sup> See, for example, *Schenck v United States*, 249 US 47 (1919) (World War I); *Dennis v United States*, 341 US 494 (1951) (communism).

<sup>63</sup> The Court held unconstitutional in these cases the actions of Southern law enforcement officers in dispersing (shall we say "prematurely") civil rights demonstrations on the ground that they would provoke a riot or other hostile audience response. See *Edwards v South Carolina*, 372 US 229, 237-38 (1963); *Cox v Louisiana*, 379 US 536, 550-51 (1965). Professors Sunstein and Fiss have interpreted these cases to require affirmative police protection of any speaker whose words arouse a threatening response. See Cass R. Sunstein, *Free Speech Now*, 59 U Chi L Rev 255, 273-74 (1992); Fiss, 100 Harv L Rev at 786 (cited in note 19). I think the decisions have a narrower meaning, consistent with the theory I have proposed. The decisions established not a duty to provide police protection for all speakers, but rather a duty to provide as much police protection for speakers whose ideas officials hate as for speakers whose ideas the officials approve.

<sup>64</sup> Ronald Dworkin has suggested another example to make a similar point. He asks why the feminist movement has focused so much attention on pornography when (by his estimation) "popular forms of [mass] culture—the view of women presented in soap operas and commercials, for example—are much greater obstacles to [ ] equality than the dirty films watched by a small minority." Ronald Dworkin, *Women and Pornography*, NY Rev Books 36 (Oct 21, 1993). He concludes that pornography, though less harmful than these other forms of culture, is more detestable—that the rawness with which it expresses the idea of sexual subordination causes it to be "deeply offensive in itself, whether or not it causes any other injustice or harm." Id. Dworkin, of course, may be wrong about the relative harms caused by these two forms of speech. But his example suggests the potential for purely "ideological" motives to influence regulatory proposals and the estimations of harm on which they are built.

<sup>65</sup> See also Schauer, *Free speech* at 82 (cited in note 39) (noting "in people a desire for

tility no doubt may derive from the fact of harm and have no significance of its own. But so too hostility toward ideas may exist apart and freestanding, or even impel the judgment of harm. Most often, perhaps, the two kinds of motives become hopelessly entangled, as one influences the other which in turn influences the first in a kind of endless feedback loop. But the complexity of this relationship—the way the different motives interact with each other, on both a conscious and an unconscious level—should not obscure the role that ideological factors may play. Hostility against speech (or sympathy toward it) may lead the government or public to overassess (or underassess) the harm speech causes. Likewise, hostility against speech (or sympathy toward it) may lead the government or public to tolerate a lesser (or greater) degree of the harm than it otherwise would. In either case, hostility (or sympathy) is doing some of the work in the decision to impose a limit on speech. The desire to suppress for its own sake—the tendency to count challenge or opposition itself as harm—is impermissibly entering into the calculation.<sup>66</sup>

The distinction between harm-based and ideological motives I have offered here differs from a distinction several other commentators have drawn relating to governmental purpose. They have argued that the great divide is between laws based on the communicative effect of speech and those based on other effects—or more narrowly, between laws based on the “persuasive” effect of speech and those based on other effects. When phrased in terms of communicative effect, the argument runs as follows. The government may not restrict speech for any reason having to do with either the messages embedded in the speech or the consequences of those messages; the government may impose restrictions for reasons relating only to aspects of the speech independent of and extraneous to the message, such as the speech’s decibel level.<sup>67</sup> When phrased, alternatively, in terms of “persuasive”

---

unanimity, an urge to suppress that with which they may disagree even if there seems no harm to that expression”).

<sup>66</sup> The complex relationship between harm and ideology has a familiar analogue in equal protection law. Discrimination on the basis of race, gender, and so forth often has a basis in reason—in accurate generalizations about the different characteristics, behaviors, and needs of members of particular groups. But such discrimination also often has a basis in fear, loathing, and prejudice. Hatred of this kind in part may emerge from actual difference, in part may exist as something independent, in part may construct and influence the perception of difference. The entanglement of hostility and harm-based reasons in First Amendment law is in many ways similar.

<sup>67</sup> The concept of communicative effects has received its fullest explication in the work of John Hart Ely and Laurence Tribe. See John Hart Ely, *Democracy and Distrust*:

effect, the argument has a narrower cast. It now posits that the government may not restrict speech for any reason having to do with either the message itself or the ability of the message to persuade listeners to take some action.<sup>68</sup> So whereas my conception of motive countenances (to use but a few examples) reasons relating to the capacity of speech to cause psychic trauma, trigger a hostile audience response, or persuade an audience to violate a law, the communicative effects theory views all of these reasons as impermissible, and the persuasive effects theory rules out the final reason, though not the two others.

These alternative theories are deficient in two respects. First, they conflate motives that I have just argued differ from each other, albeit in a complex, shifting, and elusive manner. Second, and more important for my purposes, they fail as descriptive theories of what constitutes illicit motive in First Amendment law. Courts in fact allow the government to restrict speech for reasons concededly related to its communicative (including persuasive) effects. True, the government usually must meet a heightened standard when it justifies a law on these grounds.<sup>69</sup> But if the motives identified by these theories were impermissible, in the way I use the term, a court would have to invalidate in all circumstances restrictions concededly based on them. A reason that is impermissible cannot count as a reason because it refers to a thing that cannot count as a harm. Reasons related to communicative or persuasive effects are not of this kind: the

---

*A Theory of Judicial Review* 110-11 (Harvard 1980) (asking whether "the evil the state is seeking to avert . . . arises from something other than a fear of how people will react to what the speaker is saying"); Laurence Tribe, *American Constitutional Law* § 12-2 at 789-90 (Foundation 2d ed 1988) ("[I]f the constitutional guarantee is not to be trivialized, it must mean that government cannot justify restrictions on free expression by reference to the adverse consequences of allowing certain ideas or information to enter the realm of discussion and awareness.").

<sup>68</sup> See Scanlon, 1 *Phil & Pub Aff* at 212-13 (cited in note 39); Strauss, 91 *Colum L Rev* at 334 (cited in note 39) ("[T]he government may not justify a measure restricting speech by invoking harmful consequences that are caused by the persuasiveness of the speech."). A still narrower version of this theory might posit that the government is forbidden from restricting speech on the ground that it will persuade people to adopt wrong or false opinions (rather than persuade people to take actions causing harm). If phrased in this way, the principle becomes compatible (indeed, is largely synonymous) with my description of impermissible motive.

<sup>69</sup> The heightened standard actually arises from the content-based terms of a law rather than from its underlying justification. Of course, the distinction between content-based and content-neutral laws may serve as an easily administrable device to test for impermissible motive. See Section III.A. But as I explain in the text, the structure of First Amendment law belies the view that all reasons relating to the communicative or persuasive effects of speech are impermissible.

government may count such effects—for example, violations of law arising from advocacy—as cognizable harms and may move to prevent them upon showing a real need to do so.<sup>70</sup> By contrast, what I have termed “ideological” reasons are indeed off-limits. The government may not count a challenge to governmental officials or official opinion as a harm and may not restrict speech to defeat such a challenge, even if the restriction is essential for achieving this purpose. The line between licit and illicit reasons thus lies not where these alternative theories have placed it, but between harm-based and ideological motivations.<sup>71</sup>

But if I am correct that the central prohibition of the First Amendment relates to ideological motive, then the practical import of the Amendment would seem nonexistent. Even assuming there is such a thing as a governmental motive in the sense I have used the term, how would a court ever discover the motive that I have said is off-limits? Officials will not admit (often, will not themselves know) that a regulation of speech stems from hostility or self-interest. They will invoke in each case a plausible interest, divorced from ideological disapproval, to restrict the affected expression. Then, perhaps, even a speech-protective court will have to approve the government’s action. Or will it? The next Section focuses on this question. It first discusses briefly the difficulty of making a direct inquiry into governmental motive, as well as the very coherence of this project. It then addresses, more fully, the possibility of ascertaining motive through indirect means, by using a set of rules directed to the face of legislation that will demarcate very roughly the set of governmental actions most likely to have arisen from illicit motive.

---

<sup>70</sup> The prevailing standard, emerging from *Brandenburg v Ohio*, allows the government “to forbid or proscribe advocacy of the use of force or of law violation . . . where such advocacy is directed to inciting or producing imminent lawless action and is likely to incite or produce such action.” 395 US 444, 447 (1969). In cases of standard-fare criminal solicitation, with no political or “public” character, most scholars assume a lesser standard would apply. See, for example, Kent Greenawalt, *Speech, Crime, and the Uses of Language* 110-26 (Oxford 1989). Such cases thus pose even starker counterevidence for any theory of impermissible motive based on communicative or persuasive effect. So too do cases dealing with speech proposing an illegal commercial transaction, which the government also may regulate freely under the First Amendment. See *id.* at 270-71.

<sup>71</sup> I do not claim that the notion of communicative effect has no operative meaning in the law. As I discuss in Section III, First Amendment doctrine sometimes uses this notion to aid in the search for the true impermissible—that is ideological—motive.

### C. Surmounting Problems of Proof

It has become a commonplace among both judges and scholars that the search for legislative intent—indeed, the very notion of legislative intent—raises grave problems. One set of questions relates to whether there is “a” legislative intent to be found. Consider that each legislator possesses a complex mix of hopes, expectations, beliefs, and attitudes. It is not obvious which of these mental states, or combination of them, constitutes her essential intent in voting for legislation.<sup>72</sup> Now consider that a legislature has many, perhaps hundreds of members. It is, if anything, less obvious how to combine different individual intents (assuming those exist) into a composite group purpose.<sup>73</sup>

A second set of questions assumes there is such a thing as legislative intent, but asks whether we can find it (and, if so, how). This is the aspect of the problem on which the *O'Brien* Court focused when it declared that “inquiries into congressional motives . . . are a hazardous matter.”<sup>74</sup> Often, the Court recognized, evidence of legislative purpose will consist “of what fewer than a handful of Congressmen said about it.”<sup>75</sup> But “[w]hat motivates one legislator to make a speech about a statute is not necessarily what motivates scores of others to enact it . . . .”<sup>76</sup> Indeed, what motivates one legislator to make a speech about a statute is not necessarily what motivates that very legislator to enact it.<sup>77</sup> The evidentiary materials available—floor statements,

<sup>72</sup> See Ronald Dworkin, *Law's Empire* 321-33 (Harvard 1986); *Edwards v Aguillard*, 482 US 578, 637 (1987) (Scalia dissenting).

<sup>73</sup> See Dworkin, *Law's Empire* at 320-21 (cited in note 72). The branch of public choice theory growing out of Arrow's impossibility theorem presents an especially strong challenge to the notion that a collective body has an intent. See Kenneth A. Shepsle, *Congress Is a "They," Not an "It": Legislative Intent as Oxymoron*, 12 Intl Rev L & Econ 239, 249-50 (1992); Frank H. Easterbrook, *Statutes' Domains*, 50 U Chi L Rev 533, 537-39 (1983). For a defense of the notion of legislative intent, see Andrei Marmor, *Interpretation and Legal Theory* 159-72 (Clarendon 1992).

<sup>74</sup> 391 US at 383.

<sup>75</sup> *Id.* at 384.

<sup>76</sup> *Id.*

<sup>77</sup> The Court expressed this point in a slightly different way: it worried that legislation “could be reenacted in its exact form if the same or another legislator made a ‘wiser’ speech about it.” *Id.* Congress, that is, could respond to the judicial invalidation of a statute on grounds of improper purpose by passing the identical statute with a cleaned-up legislative record. This argument, often termed the “futility” concern, usually is treated as an independent reason—distinct from the difficulty of ascertaining legislative motive—to disdain an inquiry into purpose. See Ely, 79 Yale L J at 1214-15 (cited in note 40) (accepting the argument); Paul Brest, *Palmer v. Thompson: An Approach to the Problem of Unconstitutional Legislative Motive*, 1971 S Ct Rev 95, 125-27 (rejecting the argument). The futility concern, however, rests entirely on the problem of ascertainability. If motive

committee reports, and so forth—provide a less than reliable guide to the intent of any individual legislator, let alone to the intent of the collective body.

The conception of impermissible motive I use in this Article does not fall prey to questions regarding the coherence of the notion of collective intent. The issue of motive, as I have framed it, is one of but-for causation: would the restriction on speech have passed—that is, would the outcome of the legislative process have differed—in the absence of ideological considerations?<sup>78</sup> To answer this question, it is unnecessary to consider the essential intent of any individual, much less of the decision-making body; it is irrelevant whether any such intent exists or can exist as a conceptual matter. The “thing” that a court is attempting to find is only the intrusion of a particular factor in a way that affects the decision-making process. Whatever questions attach to the notion of collective intent do not place in doubt these but-for causes of governmental action.<sup>79</sup>

But this conception of impermissible motive cannot avoid questions relating to the difficulty of finding the relevant object. True, the Court need not determine the collective sense of a decision-making body or even a single legislator’s full state of mind. But the Court must perform a task that might be as hard: determining whether a particular factor played a but-for role in a

---

could be reliably determined, the Court would not fear futility, for it then could invalidate the reenacted, no less than the original, statute (assuming the motive remained the same). The problem of futility arises only because legislators, at any time, can feign a purpose they do not have.

<sup>78</sup> Compare Laurence H. Tribe, *The Mystery of Motive, Private and Public: Some Notes Inspired by the Problems of Hate Crime and Animal Sacrifice*, 1993 S Ct Rev 1, 33 n 79 (“[T]he motive inquiry in the *O’Brien* context, for example, need go no deeper than to ask whether the law would have been enacted but for the fact that draft-card burning was being used for protest.”); Brest, 1971 S Ct Rev at 119 (cited in note 77) (“[I]t is inappropriate to ask which of several possible objectives was ‘sole’ or ‘dominant’ in the decisionmaker’s mind: an illicit motive may have been ‘subordinate’ and yet have determined the outcome of the decision.”).

<sup>79</sup> Moreover, conceptual doubts about legislative intent are irrelevant to my project, which is one of understanding the root sources of current doctrine. The Court has not allowed such doubts to prevent it from inquiring into motive—even into “sole,” “dominant,” or “essential” motive—in a variety of circumstances. For example, in determining the constitutionality of legislative action under the Establishment Clause and Commerce Clause, the Court specifically asks about legislative purpose. See *Lemon v Kurtzman*, 403 US 602, 612 (1971); *Pike v Bruce Church, Inc.*, 397 US 137, 142 (1970). And as the *O’Brien* Court itself admitted, courts routinely explore legislative motive in interpreting statutes. 391 US at 383. Even if there is no such thing as legislative intent, the Court often acts as if there is. So long as this is true, objections to the concept of legislative intent do no damage to the claim that some aspect of doctrine, explicitly or implicitly, attempts to discover the intent of the legislature.

decision-making process. This task, in its most simplified form, involves reckoning how many legislators the impermissible consideration swayed and comparing that number to the margin of victory.<sup>80</sup> What *O'Brien* said about the hazards of inquiring into motive seems to apply in full to this inquiry. The standard evidence of legislative process provides an insufficient basis to make the requisite head count or even estimate its outcome.

Consider first, to highlight the difficulty of the endeavor, how a court would decide whether improper motive tainted the decision of even one legislator. Few legislators, of course, will admit to a constitutionally illegitimate purpose; the legislator instead will point to some real harm that the speech causes. Such a harm-based pretext usually will be available; cases in which the government tries to curtail speech that cannot plausibly be described as harmful are not common. Further, the legislator herself often will not know whether an illicit reason tainted her consideration of the law at issue, given the complex dynamic between legitimate assessments of harm and illegitimate attitudes toward opinions. To make matters worse, the judge handling the matter will possess her own views of the ideas restricted, which may affect her evaluation of the legislator's motives in the same diffuse and incalculable ways as the legislator's views initially affected her decision. Now consider how these difficulties multiply when a court must face the issues of aggregation involved in determining how illicit motive affected a multimember body. Hence the message of *O'Brien*: direct inquiry into motives for restricting speech very rarely will prove productive.

But the impracticality of this inquiry need not force courts to abandon the goal of invalidating improperly motivated legislation, if they can find another, more feasible way of pursuing that project. If courts cannot determine motive directly, by exploring

---

<sup>80</sup> There are many complications in determining whether ideological factors altered a legislative decision that this simplified statement of the problem ignores. Most notably, this statement overlooks the disproportionate influence that some legislators wield because of, among other things, their agenda-setting ability or their strategic importance on other issues. What if, for example, only one legislator harbored impermissible reasons for favoring a statute (all other legislators having legitimate, harm-based reasons), but she was the person responsible for bringing the statute to a vote? Or what if other legislators acceded to her wishes on the statute (having no strong views of their own) to get her vote on another issue? The critical issue is whether impermissible reasons altered the outcome of the decision-making process; in both these cases, they did. But in both cases, it will not suffice to ask whether improper motives directly accounted for the votes necessary to enact the statute. The story about how improper motive affected the outcome is more complicated, raising even greater problems of discovery.

what went into the legislative process, perhaps they can determine motive obliquely, by looking at what came out of it. Suppose that courts could develop rules relating only to the terms of legislation; suppose further that these rules predictably operated to sort out actions that had impermissible motives from those that did not. By using these rules, courts could invalidate laws supported by improper reasons without ever confronting the problems of proof generated by a direct inquiry into motive. The function of the rules in flushing out impermissibly motivated actions might not be articulated or even understood. The rules would operate in an autonomous manner, removed from explicit consideration or discussion of the question of motive. But the two would remain integrally connected: the concern with motive would determine the scope of the rules, and the rules would give effect to the concern with motive.

So it might be in First Amendment law: perhaps the Supreme Court has constructed a set of rules that allows a judge to ferret out impermissible motives at the same time as it obviates any need to ask about this issue. We might think of these rules as proxies for a direct inquiry into motive or as rules of an evidentiary nature. These rules use objective criteria, focusing on what a law includes and excludes, on what classifications it uses, on how it is written. But in making such inquiries, the rules in fact serve as an arbiter of motive. Through use of these objective tests and rules, some rough sorting out takes place: between laws tainted by ideological motives and those not so blemished.

The roughness of this division should not be understated: these tests of governmental purpose necessarily will be imperfect—simultaneously under- and overinclusive. Still, it makes sense to use the rules, rather than to ask directly about motive, as it often makes sense to use rules rather than to rely on their underlying reasons. If courts cannot reliably (or cannot at all) determine whether the reason for the rules (here, improper motive) exists, then the mistakes made without any rules will exceed the mistakes arising from the rules' structure.<sup>81</sup> The decision to use such rules thus follows from the combination of one fundamental principle and one unfortunate fact. The principle is

---

<sup>81</sup> For general discussion of this rationale for the use of rules, see Frederick Schauer, *The Second-Best First Amendment*, 31 *Wm & Mary L Rev* 1, 9-12 (1989). Just as rules better enable courts to determine legislative motive, the rules may perform the same function for legislators. Odd as it may seem, it may be easier for legislators to follow rules relating to the terms of a law than to follow a command not to consider illicit factors.

that the First Amendment bans restrictions on speech arising from hostility, sympathy, or self-interest. The fact is that courts cannot enforce this ban directly.

If all this is so, the First Amendment rules of which I am speaking, though seemingly substantive in content, resemble in function such procedural mechanisms as presumptions and shifting burdens of proof. Consider how a different body of law responds to the difficulty of proving motive. Under the labor laws, an employer may not discharge an employee because of union activity. A court deciding whether such an act has occurred will shift the burden of proof on the question of motive to the employer once the employee has made a lesser (*prima facie*) showing.<sup>82</sup> In so doing, the law in effect establishes a rebuttable presumption: the law presumes improper motive from a set of facts merely suggestive of it unless the employer proves its absence. The rules of First Amendment law work in a similar manner. They too operate, though not overtly, to make a rebuttable determination of improper motive on the basis of some set of facts—for example, a content-based classification—suggestive but not dispositive of it. It is in this sense that I have spoken of these rules as evidentiary in nature: they, no less than such procedural mechanisms as presumptions and shifting burdens, serve to ameliorate troublesome problems of proving motive by giving exceptional weight to certain evidentiary materials.

This hypothesis suggests a reinterpretation of *O'Brien*. No longer should that decision be viewed as a broad-scale stricture against invalidating regulations of speech on the basis of improper motive. That understanding of the case has always conflicted with too much in the Court's rhetoric and decisions. *O'Brien* stands for a narrower proposition, relating not to the propriety of inquiring into motive, but to the means by which to conduct this inquiry. To be more precise, *O'Brien* stated not that motive was irrelevant, but only that it could not be proved by traditional methods. In so doing, the decision left open the option of adopting a different mechanism to discover motive. The Court, as the next Section of this Article shows, has chosen this course in its elaboration of First Amendment doctrine.

---

<sup>82</sup> See *NLRB v Transportation Management Corp*, 462 US 393, 403 (1983). A similar though less potent procedural mechanism, designed to accomplish the same object, is used in Title VII cases. See *Texas Department of Community Affairs v Burdine*, 450 US 248, 252-53 (1981).

### III. THE DOCTRINE OF IMPERMISSIBLE MOTIVE

Let us accept that the First Amendment prohibits restrictions on speech stemming, even in part, from hostility, sympathy, or self-interest. And let us accept that the difficulty of proving this impermissible motive—resulting, most notably, from the government's ability to invoke pretextual reasons—gives rise to a set of rules able to flush out bad motives without directly asking about them. What would these rules look like?

The first rule would draw a sharp divide between content-based and content-neutral restrictions, with a fuzzier line bisecting the world of content-based restrictions into those based on viewpoint and those on subject matter. The second and third rules would specify exceptions to the first: instances in which a restriction, though content-neutral, demands heightened scrutiny because of suspect origin; instances in which a restriction, though content-based, could receive relaxed scrutiny because apparently safe. And the fourth rule would draw another sharp distinction, this time between actions directly addressed to speech and those affecting speech only incidentally.

These rules—the rules that would be devised to flush out illicit purpose—in fact constitute the foundation stones of First Amendment doctrine. Examining their structure reveals that the search for impermissible motive animates the doctrine, as the doctrine implements the search for motive.

#### A. The Distinction Between Content-Based and Content-Neutral Laws

The distinction between content-based and content-neutral regulations of speech serves as the keystone of First Amendment law.<sup>83</sup> Content-neutral restrictions on speech—restrictions that by their terms limit expression without regard to what is said—usually are subject to a fairly loose balancing test. So, for example, in reviewing a law that bans all billboards within city limits, the Court might consider the strength of the state interests asserted (say, in aesthetics and traffic safety), the availability of alternative means to protect those interests, the extent to which the law limits expression, and the existence of alternative avenues of communication. This analysis may well result in a decision that the law accords with the First Amendment. Con-

---

<sup>83</sup> The fullest description and analysis of this distinction remains Stone, 25 *Wm & Mary L. Rev* at 189 (cited in note 15).

tent-based restrictions on speech—restrictions that by their terms limit expression on the basis of what is said—usually are subject to far more rigorous scrutiny. This is true even in cases like *R.A.V.* in which the government concededly could restrict the speech affected through a broader law written in content-neutral language. Formulations of the standard used to review content-based action vary, but the Court most often requires the government to show a compelling interest that could not be attained through less restrictive means. Application of this standard usually leads to a law's invalidation.

Somewhat mitigating the starkness of this scheme, a further, hazier distinction operates within the realm of content-based regulation. Here, the Court often differentiates between viewpoint-based restrictions and all other content-based restrictions, including, most notably, restrictions based on subject matter.<sup>84</sup> So, for example, the Court would treat differently a law prohibiting the use of billboards for all political advertisements and a law prohibiting the use of billboards for political advertisements supporting Democrats. The former might meet constitutional standards; the latter would never succeed in doing so. It is not so much that the Court formally uses two different standards for subject matter and viewpoint regulation; in most contexts, a strict scrutiny standard applies to content-based action of all kinds.<sup>85</sup> But the Court, when reviewing subject-matter restrictions, either may apply a purportedly strict standard less than strictly or may disdain to recognize the law as content based at all.<sup>86</sup> By contrast, the Court almost always rigorously reviews and then invalidates regulations based on viewpoint.

This scheme makes no sense under the speaker-based model of the First Amendment.<sup>87</sup> Recall that this model treats as critical the sum total of expressive opportunities; the more a law

---

<sup>81</sup> See generally Geoffrey R. Stone, *Restrictions of Speech Because of its Content: The Peculiar Case of Subject-Matter Restrictions*, 46 U Chi L Rev 81 (1978).

<sup>85</sup> In some rare contexts—most notably, in nonpublic forums—the Court explicitly adopts different standards for subject-matter and viewpoint regulation. See Kagan, 1992 S Ct Rev at 42-43 (cited in note 14).

<sup>86</sup> For a case in which the Court applied a toothless version of strict scrutiny, see *Burson v Freeman*, 504 US 191, 198-211 (1992). For a case in which the Court pretended that a subject-matter restriction was content neutral, see *Rowan v United States Post Office Department*, 397 US 728, 737-38 (1970). Of course, in many subject-matter cases, the Court applies a strict scrutiny standard with all the rigor its name implies. See, for example, *Police Department of Chicago v Mosley*, 408 US 92, 98-102 (1972).

<sup>87</sup> See Redish, 34 Stan L Rev at 128-39 (cited in note 34) (criticizing the distinction on this ground); Stone, 25 Wm & Mary L Rev at 197 (cited in note 15) (acknowledging the point, but approving the distinction).

curtails the ability to speak, the greater its constitutional difficulty. Yet a content-neutral law, no less than a content-based law, can lessen the ability to speak; indeed, a content-neutral law can do so more dramatically. To use my earlier example, a general ban on billboards will reduce speech more than a ban on billboards for political advertisements, which in turn will reduce speech more than a ban on billboards disabling only Democrats. Yet under current law, the Court will subject the first of these ordinances to the most relaxed form of review and the last to the strictest. Consider in the same vein the cases of *Police Department of Chicago v Mosley*<sup>88</sup> and *Grayned v City of Rockford*.<sup>89</sup> In the latter, the Court upheld a content-neutral ban on speech in the vicinity of a school; in the former, the Court invalidated a similar ban on the ground that it exempted speech about labor disputes from its general prohibition. Finally, recall the Court's view in *R.A.V.* that although a ban on all fighting words would have passed muster, with the category of fighting words treated as content neutral,<sup>90</sup> a ban on fighting words limited to a certain subject—worse, to a certain viewpoint—violated constitutional norms. Each of these examples shows that a concern with the extent of expressive opportunities cannot explain the most critical aspect of First Amendment doctrine.

Perhaps, however, a concern with the quality of the speech market—the concern of the audience-based model—may explain the distinction between content-based and content-neutral regulation.<sup>91</sup> The argument, anticipated in my discussion of *R.A.V.*, relies on the “distorting” effect of content-based, and especially viewpoint-based, regulation. The edict “no billboards” on its face handicaps equally all ideas. The edict “no ads for Democrats on billboards,” by contrast, disadvantages certain ideas to the benefit of others. Finally, the edict “no political ads on billboards” falls in between these extremes. Bans of this kind at the least disfavor one subject of discussion compared with others. And they often (at any rate, more often than content-neutral restrictions) oper-

---

<sup>88</sup> 408 US 92 (1972).

<sup>89</sup> 408 US 104 (1972).

<sup>90</sup> The Court analogized the regulation of fighting words to the regulation of sound trucks, which of course is content neutral. The Court explained: “[F]or purposes of [the First] Amendment, the unprotected features of [fighting] words are, despite their verbal character, essentially a ‘nonspeech’ element of communication.” *R.A.V.*, 505 US at 386. I discuss later the reasons for treating the category of fighting words as neutral with respect to the content of speech. See text accompanying note 182.

<sup>91</sup> For such an argument, see Stone, 25 *Wm & Mary L Rev* at 217-27 (cited in note 15); Sunstein, *Democracy and the Problem of Free Speech* at 170 (cited in note 19).

ate to skew debate among competing ideas on a single subject: consider, for example, if in 1970 the government had banned discussion of the Vietnam War. Perhaps, then, current doctrine responds to the different ways in which viewpoint-based, subject-matter-based, and content-neutral laws distort public discourse and thereby (in Meiklejohn's phrase) mutilate the community's thinking process.

Even assuming, however, that a law disparately affecting ideas necessarily skews the speech market—an assumption I contest shortly—this justification of the divide between content-based and content-neutral regulation suffers from two related weaknesses.<sup>92</sup> First, a doctrinal structure based on the problem of distortion seemingly would subject to heightened scrutiny whatever content-neutral rules fall much more heavily on one idea than others. Suppose, for example, that only Democrats, and not Republicans, use billboards to advertise; then, the skewing effect of a general ban on billboards would match the skewing effect of a law specifically barring Democrats from this forum. To put the point more generally, content-neutral laws often have content-based effects—and sometimes these are quite dramatic. A jurisprudence concerned with distortion should treat these cases with the utmost seriousness. But current doctrine all but ignores the distorting effects of content-neutral law.<sup>93</sup>

Second and conversely, a body of doctrine based on the problem of distortion apparently would subject to relaxed review any content-based laws that have only a modest tilting effect. Consider again our viewpoint-based billboard regulation; if neither Democrats nor Republicans use billboards, disallowing such use for one party only will not skew public discourse. Or recall again *R.A.V.* where the ban on racist fighting words could not seriously have distorted the deliberative process. In such cases, the small quantity of speech affected, combined with the ready availability of alternative means to communicate the “handicapped” idea, makes the danger of distortion insignificant. Yet First Amendment doctrine distinguishes not at all between content-based

---

<sup>92</sup> I consider below whether a model focusing on motive also suffers from these weaknesses and conclude that it does not—or, at least, that it does not do so to the same extent. See text accompanying note 104.

<sup>93</sup> For discussion of the cases, see Stone, 25 *Wm & Mary L Rev* at 218-22 (cited in note 15). For an argument that content-neutral laws with significant content-based effects ought to be treated as if facially content-based because of the extent to which these laws distort debate, see Susan H. Williams, *Content Discrimination and the First Amendment*, 139 *U Pa L Rev* 615, 655-63 (1991).

laws of this kind and content-based laws that wholly excise ideas from public discourse.<sup>94</sup>

One explanation of these oddities refers to the difficulty of deciding when a regulation has a skewing effect that is sufficiently large or small to alter the usual standard of review.<sup>95</sup> Any such decision necessarily will involve difficult questions relating to what speakers use what forms of speech, as well as to how effectively different forms of speech (both the form restricted and its alternatives) communicate a desired message. Perhaps, given these difficulties, the distinction between content-based and content-neutral regulation functions as an imprecise (both over- and underinclusive) but still sensible mechanism for sorting out consequential from inconsequential skewing effects, in accord with the dictates of the audience-based model.

This explanation, however, is unconvincing. True, there would be hard cases if courts evaluated skewing effects—cases involving tricky issues of measurement and line drawing. But such problems seem no more common or intractable in this adjudicative context than in many others, where no one thinks they preclude evaluative efforts. There also would be many simple cases as measured by an effects-based ruler—cases in which courts confidently could say either that a content-based law would have minor skewing effects or that a content-neutral law would cause major distortion. Recall again *R.A.V.*, where the ordinance would have had utterly insignificant skewing effects. Were courts primarily concerned with distortion they would at least modify the strict distinction between content-based and content-neutral laws to respond to the host of cases in which they could directly evaluate skewing effects.

Indeed, to the extent this conclusion is wrong, it is so because of a fear of improper motive. Suppose, that is, there is some special reason to resist case-by-case line drawing with respect to the skewing effects of a speech restriction. What would this reason be? It likely would relate to the fear that a judge's own biases toward the speech affected would taint her decision as to whether the restriction had a severe or narrow skewing effect.<sup>96</sup> But if this is the reason for preventing judges from mak-

---

<sup>94</sup> For discussion of the case law, see Stone, 25 *Wm & Mary L Rev* at 200-01 (cited in note 15).

<sup>95</sup> See *id* at 224-27.

<sup>96</sup> See *id* at 225 (Judges "may be influenced by . . . biases that may undermine their ability to evaluate accurately and impartially the extent to which particular . . . restrictions actually impair the communication of specific, often disfavored, mes-

ing case-specific determinations as to distortion, then the doctrine arises from considerations of motive at least to this extent: that fear of illicit motive constrains and structures the inquiry as to effects. And if this is so, then it is at least true that the doctrine attempts not to create a distortion-free universe, but only to accomplish as much as can be accomplished in this direction consistent with an omnipresent fear of improper motive.

But much more than this can be said. The discussion so far has assumed that the disparate impact of a law on ideas will distort the speech market. If that assumption is false, then the distinction between content-based and content-neutral laws—even if the most sensible way of determining whether a law disparately affects ideas—would not further the interest in balanced discourse.

In fact, this assumption is hard to defend, for as a conceptual matter, the disparate impact of a law on a set of ideas might lead to balance as easily as to distortion.<sup>97</sup> Remember that each regulation affecting speech acts against a backdrop of countless other regulations affecting speech, sometimes directly, sometimes incidentally. Among these are rules of property and contract, which provide some speakers with access to the most effective means of expression and consign other speakers to the least so. All these regulations, operating together, give shape and content to the realm of discourse, and given the nature of these rules—specifically, the ways they effect inequalities of wealth and access—the speech environment they create stands little chance of nearing the ideal condition. Distortion, skew, tilt—whatever one calls lapses from the ideal—will occur all over. In such a setting, any law with a disparate impact on ideas may succeed in balancing, no less than skewing, the speech market; conversely, any law affecting ideas equally may perpetuate a skewed, no less than a balanced, speech environment. As a logical matter, such laws will do the one thing no less than the other.

The distinction between content-based and content-neutral regulation thus cannot rest on a pure audience-based approach to the First Amendment. It is true that this distinction reflects the likelihood that a law will change the prevailing structure of pub-

---

sages.”); Ely, *Democracy and Distrust* at 112 (cited in note 67) (Such evaluations “inevitably become involved with the ideological predispositions of those doing the evaluating.”).

<sup>97</sup> See text accompanying notes 19-20. For fuller discussion, see Owen M. Fiss, *Free Speech and Social Structure*, 71 *Iowa L. Rev.* 1405, 1410-13 (1986); Sunstein, *Democracy and the Problem of Free Speech* at 177-80 (cited in note 19).

lic comment. But if that structure itself departs from the ideal—if, in Meiklejohn's words, the existing distribution of views itself "mutilat[es]" the "thinking process of the community"<sup>98</sup>—then the presumption against content-based law may not serve to protect this process. In such a world (which is our world), a content distinction has no necessary tendency to impede the goals of the audience-based model.

Indeed, this model of the First Amendment might well command (not merely tolerate) the use of content discrimination in some circumstances. As one proponent of the view has urged, "governmental action . . . based on content . . . might be needed to protect our freedom" by ensuring that "public debate [is] enriched and our capacity for collective self-determination enhanced."<sup>99</sup> No proposal could be further from current doctrine; the use of the audience-based model seems to counsel discarding the keystone of the law for its opposite.

Few courts or commentators would view with equanimity a reform of this kind; indeed, even proponents of the audience-based model might favor the Court's decision to use as its benchmark the actual, rather than the perfect, distribution of viewpoints. The stated reason might run something as follows: no matter how unhealthy the existing speech market, governmental action directed at the content of speech would cause in most cases further harm. But given all I have said, how is it possible to defend this assertion? The answer to this question will suggest the deepest wellspring of First Amendment doctrine—the concern that drives and indeed defines all others.

One explanation for our choice of benchmark refers to the difficulty—even the incoherence—of defining the ideal realm of discourse.<sup>100</sup> We do not have a full picture of what a well functioning marketplace of ideas would look like. Who would say what in such a system? At what point would an idea become over- or underrepresented? Perhaps we can provide no account of the optimal mix of expression. And if we cannot describe the ideal, perhaps we also cannot decide whether an action would bring us closer to, or take us further from, this state of perfec-

---

<sup>98</sup> Meiklejohn, *Political Freedom* at 27 (cited in note 15) (emphasis omitted).

<sup>99</sup> Fiss, 71 *Iowa L. Rev.* at 1415 (cited in note 97). See also Sunstein, 59 *U. Chi. L. Rev.* at 290-91 (cited in note 63) ("[E]fforts to restructure the marketplace [of ideas] might even be seen as the discharge of the legislature's constitutional duty, a duty that courts are reluctant, for good institutional reasons, fully to enforce.").

<sup>100</sup> For fuller discussion of these issues, see Strauss, 91 *Colum. L. Rev.* at 349 (cited in note 39); Strauss, 1993 *U. Chi. Legal F.* at 202-10 (cited in note 20).

striction poses a greater risk of improper purpose than the usual content-neutral law. Consider again a ban on political speech in a given forum (in my running hypothetical, on billboards). Though neutral reasons may support such a ban,<sup>103</sup> there is heightened danger that the government is acting in part for illicit reasons—because, say, officials know that a disfavored political group disproportionately uses this mode of communication. Such a purpose could infect as well the law banning all billboards, but as the law applies to an ever greater range of ideas, the probability of taint decreases.

Finally, the focus on purpose explains why First Amendment doctrine ignores the severe skewing effects of some content-neutral laws and the slight skewing effects of some viewpoint-based laws. It is true that such effects may offer evidence relating to motive; presumably the greater the skewing effect, the greater the chance that illicit considerations have intruded. But in most cases a law's terms more reliably indicate illicit motive than its effects and thus should control the legal analysis.<sup>104</sup> A content-neutral law, even when it has severe skewing effects, poses only a minor risk of improper motive because the law creates such effects along so many dimensions. The diffuseness of the law outweighs the severity of its impact on any particular idea as evidence of motive. Conversely, a content-based law, even when it has insignificant skewing effects, presents a substantial risk of impermissible motive because the effects occur in so narrow an area. The focused nature of the law outweighs the mildness of its impact on an idea as evidence of motive. This is why courts would treat differently our three billboard laws even if, as could be true, the laws similarly affected the distribution of political views. The terms of the laws indicate, even if the effects of the laws do not, disparate risks of improper motive.

The distinctions among viewpoint-based laws, other content-based laws, and content-neutral laws thus create a set of pre-

---

<sup>103</sup> The government offered such legitimate justifications in *Greer v Spock*, 424 US 828 (1976), and *Lehman v City of Shaker Heights*, 418 US 298 (1974), which involved bans on political speech on an army base and in a mass transit system's advertising space. The government in the former case invoked the interest in keeping the military removed, in both appearance and reality, from partisan causes. *Greer*, 424 US at 839. The government in the latter case pointed to administrative problems involved in allocating limited space to political candidates. *Lehman*, 418 US at 304. Both of these motives, in the scheme of this Article, are perfectly permissible.

<sup>104</sup> I here put to one side the concern that bias will infect the measurement of a law's skewing effects. These concerns only strengthen the conclusion that the face of a law indicates more reliably than the effects of the law what purposes underlie it.

sumptive conclusions about when improper motive has tainted a restriction on speech. These distinctions will suggest some "wrong" results, for some viewpoint-based laws arise solely from legitimate reasons and some content-neutral laws arise partly from their opposite. We tolerate the imprecision because the alternative—a direct inquiry into motive—will produce even more frequent errors. If the facial markers we use are not perfect, they are better than what they replace.

We also attempt to mitigate the imprecision by making the outcomes suggested by these facial distinctions presumptive only and requiring courts to consider evidence to the contrary. But this evidence of motive again takes an indirect form: it resides at this stage in the substantiality of the asserted legitimate interest for the restriction and the closeness of the fit between that interest and the terms of the law.

So, for example, the strict scrutiny standard—indeed, each component of it—is best understood as an evidentiary device that allows the government to disprove the implication of improper motive arising from the content-based terms of a law.<sup>105</sup> This is true first of the compelling interest requirement: the stronger the state interest asserted, the more likely it is that the government would act to achieve that interest in the absence of antipathy toward the speech affected. Similar reasoning applies to the demand for close tailoring. If a restriction applies to more speech than necessary to achieve the interest asserted, the suspicion deepens that the government is attempting to quash ideas as ideas rather than to promote a legitimate interest. And if a restriction applies to less speech than implicates the asserted interest, so too the concern grows that the interest asserted is a pretext.<sup>106</sup> But if a restriction fits along both dimensions—if it applies to all and also to only the speech that threatens the assert-

---

<sup>105</sup> John Hart Ely has made a similar argument on the use of strict scrutiny in equal protection law. See Ely, *Democracy and Distrust* at 145-48 (cited in note 67). There, a compelling interest and a close fit between means and end negate the presumption of illicit motive arising from use of a suspect classification such as race. In the First Amendment context, matters are the same, except that the suspect classification is content.

<sup>106</sup> In effect, the content-based nature of the law has raised suspicions so great that the usual defense of taking "one step at a time," see *Williamson v Lee Optical Co.*, 348 US 483, 489 (1955), is unacceptable. We view the underinclusive action not as a first step toward achieving a legitimate end, but as confirmation of an illegitimate purpose. See, for example, *Erznoznik v City of Jacksonville*, 422 US 205, 214-15 (1975) (holding that an ordinance prohibiting nudity in drive-in movie theaters could not be justified as a traffic regulation because "a wide variety of other scenes in the customary screen diet . . . would be [no] less distracting").

ed interest—then there is an assurance that the government has acted for proper reasons. In this way, the strict scrutiny test operates as a measure of governmental motive. The showing that the government must make under that standard does not serve, as on a scale, to outweigh impermissible motive or counter its harms. The showing instead serves an evidentiary function: to disprove (again, of necessity indirectly) the inference of bad motive that arises from the content-based face of a law.<sup>107</sup>

A similar mechanism, though operating in reverse, is necessary to assess the constitutionality of content-neutral laws. So long as a content-neutral law has differential effects on particular ideas—even assuming those effects are widely dispersed—it may bear the taint of improper motive. Officials may care so much about suppressing a particular idea affected by a content-neutral law as to disregard or tolerate the law's other consequences. Or, in a slightly different vein, officials may desire a broad-scale entrenchment of status quo positions and enact a law restricting all expression in a certain sphere in order to achieve this object.<sup>108</sup> Such a restriction, in addition to benefiting ideas already accepted, allows the government to emerge as the dominant speaker in the sphere, able to control opinion through speaking itself rather than through regulating the speech of private parties. For these reasons, the presumption that content-neutral laws are untainted by impermissible motive must remain just that—a presumption subject to rebuttal.

The relatively deferential standard governing the constitutionality of content-neutral laws serves this kind of evidentiary function. Just as passing strict scrutiny demonstrates that a content-based law has a legitimate purpose, thus rebutting a presumption of impropriety, so flunking this looser standard demonstrates that a content-neutral law has an illegitimate basis, thus rebutting the opposite presumption. The way this occurs should by now be clear. The less significant the legitimate inter-

---

<sup>107</sup> Justice Kennedy has expressed a similar view of the function of the strict scrutiny standard in First Amendment law. He has written that "the compelling interest test . . . determin[es] the accuracy of the justification the State gives for its law." *Freeman*, 504 US at 213 (Kennedy concurring). I take this to be another way of saying that the strict scrutiny standard is a tool for discovering the government's real motive.

<sup>108</sup> Subject-matter restrictions accomplish a narrower de facto favoring of status quo positions: for example, a law prohibiting all discussion of the government's foreign policy will favor whatever views on that policy are currently dominant. In such a case, the danger of impermissible motive is significant because of the restricted scope of the law (and of the consequent skewing effect). Content-neutral laws, which favor status quo positions generally, pose a lesser, but still cognizable, danger.

est supporting the law, the greater the reason to distrust the government's action. Similarly, the looser the fit between the interest asserted and the contours of the law, the greater the cause for suspicion. At a certain point—when the asserted interest is insubstantial or when it does not fit the scope of the challenged regulation—the usual presumption of proper purpose topples; there is reason, then, to think that the law, though content neutral, has been tainted by impermissible purpose.<sup>109</sup>

I do not mean to say that review of content-neutral regulations serves only, or even primarily, as a mechanism to discover bad motive. The review of content-neutral laws also functions to ensure adequate expressive opportunities, in keeping with what I have called the speaker-based perspective. If expressive activity has special value to individuals, then the government should have to justify in a special way, by offering unusually weighty countervailing interests, any restriction on expression.<sup>110</sup> In this way, one of the effects-based models supplements the purpose-based model in explaining the standard of review for content-neutral legislation.<sup>111</sup>

---

<sup>109</sup> The Court's "public forum" doctrine can be seen as establishing a kind of backstop to this relatively easy test for illicit motive. Even if a content-neutral law tainted with improper motive manages to pass this test, the public forum doctrine may prevent the law from operating in certain places. Thus, public forums serve as a kind of safe harbor for speech, providing in certain areas an extra level of protection. Of course, public forum doctrine also functions to ensure a minimum level of opportunities for expression, in line with the speaker-based model of the First Amendment. Viewed from this perspective, public forum doctrine insulates a zone of speech against even properly motivated governmental action; only outside that restricted zone does the question become one of motive.

<sup>110</sup> This rationale explains the tendency of the Court, in practice, to review content-neutral regulations more strictly when they have a severe effect on expressive opportunities and less strictly when they have a modest effect. See Geoffrey R. Stone, *Content-Neutral Restrictions*, 54 U Chi L Rev 46, 58-59 (1987). Under the speaker-based theory, as the effect of a law on expressive opportunities increases, so too should the government's burden of justification. A motive-based theory might add a further reason for varying the standard of review in this way. As the impairment of expressive opportunities increases relative to the importance of the asserted governmental interest, so too does the suspicion grow that the government is acting for illegitimate reasons.

<sup>111</sup> The audience-based model of the First Amendment explains less well than its competitors the standard of review applied to content-neutral legislation. If "[w]hat is essential is not that everyone shall speak, but that everything worth saying shall be said," Meiklejohn, *Political Freedom* at 26 (cited in note 15), then some content-neutral restrictions on speech (like some content-based restrictions) will count not as harms, but as positive goods. The "traditional American town meeting," *id* at 24, which Meiklejohn used as the model of public debate, indeed depends on rules of order. Some content-neutral laws, of course, may restrict speech so broadly as to disserve the interests of the audience. But the audience-based theory, taken alone, cannot explain the practice of subjecting not only these but all content-neutral laws to special scrutiny. For that result to follow, it is necessary to refer either to the interests of potential speakers or to the possibility of

But only the purpose-based model can explain the difference in the levels of review applicable to content-based and content-neutral laws. In justifying this distinction, the speaker-based model provides no assistance. The audience-based model becomes useful only to the extent that it reduces to a motive inquiry, concerned not with skewing effects per se, but only with such effects as arise from, and are defined by, illicit purpose. It is the quest for purpose that drives the doctrine—that explains the divergent treatment given to content-based and content-neutral laws. So too that quest explains the creation of exceptions to this doctrinal division: instances in which the courts treat facially content-neutral laws as content based or facially content-based laws as content neutral. The next two Sections deal with these exceptions and the role of motive analysis in explaining them.

#### B. "Suspect" Content-Neutral Laws

Although most content-neutral laws pose small danger of stemming from improper purpose, some such laws present greater risk. Three kinds of content-neutral laws, in particular, raise the same specter of improper purpose as the typical content-based governmental action. This suspect trio of facially content-neutral laws consists of (1) laws conferring standardless discretion on administrative officials; (2) laws turning on the communicative effect of speech; and (3) laws attempting to "equalize" the speech market. Because these laws, though content neutral, sound the alarm of illicit purpose, First Amendment doctrine treats them just as it does content-based actions. Here too, the standard applied operates as a presumption of improper motive, adopted in response to difficulties of proof.

##### 1. Laws conferring discretion.

Consider this law: "It shall be unlawful for any person to maintain and operate . . . any radio device, mechanical device, or loud speaker . . . which is so placed and operated that the sounds coming therefrom can be heard to the annoyance or inconvenience of travelers upon any street . . . [except if] the same be done under permission obtained from the Chief of Police."<sup>112</sup>

In *Saia v New York*, the Court held this ordinance unconstitutional on its face because it prescribed "no standards . . . for

---

improper purpose on the part of legislators or judges.

<sup>112</sup> *Saia v New York*, 334 US 558, 558-59 n 1 (1948).

the exercise of [the police chief's] discretion."<sup>113</sup> The Court likewise has invalidated laws broadly empowering officials to grant (or deny) permission to distribute leaflets, hold parades, or erect newsracks on public property—all, again, on the ground that the laws failed to cabin appropriately administrative discretion.<sup>114</sup> In none of these cases did the Court wait to find that the administrator actually had abused her discretion. In several of the cases, the administrator had no opportunity to engage in abuse because the speaker never applied for the requisite license.<sup>115</sup> In most of the cases, the administrator might well have succeeded in denying the license (assuming the speaker had applied) under a properly drafted statute. The constitutional problem in the cases thus arose not from any administrative decision restricting speech, but from the wide authority that the statute gave to administrators to restrict speech, both then and in the future.<sup>116</sup> Otherwise put, the rule established in these cases responded not to any actual abuse of governmental authority, but only to the potential for abusive conduct.

What could account for a rule of this sort? The best explanation, stated briefly for now, goes as follows. A law conferring standardless discretion effectively delegates to administrators the power to make decisions about speech on the basis of content. Such administrative decisions raise the same constitutional concern as do content-based laws: the danger of improper motive. Because courts cannot discover improper purpose directly, they use as a proxy the presumption against content-based distinctions. But likewise, because courts cannot easily determine, in the context of administrative action, when a content-based decision has occurred, they here add a further prophylactic rule, designed both to prevent and to detect content-based administration. The fundamental purpose of this rule barring standardless discretion thus resides in its capacity to assist in the campaign against impermissible motive.<sup>117</sup>

---

<sup>113</sup> 334 US 558, 560 (1948).

<sup>114</sup> See *Lovell v City of Griffin*, 303 US 444, 450-51 (1938) (leaflets); *Shuttlesworth v City of Birmingham*, 394 US 147, 150-51 (1969) (parades); *City of Lakewood v Plain Dealer Publishing Co.*, 486 US 750, 769-70 (1988) (newsracks).

<sup>115</sup> See *Lovell*, 303 US at 448; *Lakewood*, 486 US at 754.

<sup>116</sup> As the Court stated in *Freedman v Maryland*, "[i]n the area of freedom of expression . . . one has standing to challenge a statute on the ground that it delegates overly broad licensing discretion to an administrative office, whether or not his conduct could be proscribed by a properly drawn statute, and whether or not he applied for a license." 380 US 51, 56 (1965).

<sup>117</sup> A similar argument explains at least in part the First Amendment's vagueness

To amplify this understanding of the rule, it is best to begin by rejecting the alternatives. The speaker-based model cannot explain the rule against standardless licensing schemes because such schemes do not necessarily curtail more speech than other, less constitutionally suspect modes of restricting expression. To see this point, contrast *Saia* to *Kovacs v Cooper*,<sup>118</sup> decided seven months later. In *Kovacs*, the Court upheld an ordinance banning the use “upon the public streets” of any “sound truck, loud speaker, or sound amplifier.”<sup>119</sup> The amount of speech that this ordinance limited must have exceeded the amount that the law in *Saia* curtailed; the latter, after all, allowed exemptions from the ban. The constitutional difficulty in *Saia* thus could not have stemmed from the extent of the restriction. The same is true of all standardless licensing schemes, which may restrict less, as well as more, speech than a cabined licensing system or a flat prohibition.<sup>120</sup> Nothing in the nature of standardless licensing indicates the amount of speech either actually or potentially restricted. If this is true, nothing in the nature of standardless licensing makes it peculiarly problematic from the standpoint of maximizing communicative outlets.

Similarly, the audience-based model fails to explain the rule against standardless licensing schemes. Standardless licensing, as I will show, may well encourage content-based decisions, which punish and deter certain ideas as compared to others.<sup>121</sup>

---

doctrine. Vague laws delegate to administering officials a kind of standardless discretionary authority—here, to interpret and apply unclear language. Such discretion raises much the same concerns of improper motive as I discuss in the text.

<sup>118</sup> 336 US 77 (1949).

<sup>119</sup> *Id.* at 78.

<sup>120</sup> Another example comes from *Lakewood*, where the Court invalidated an ordinance regulating the placement of newspaper vending machines for failure to limit properly the relevant official’s discretion. 486 US at 750. All Justices assumed for purposes of the case that the city could have chosen to ban all such vending machines. *Id.* at 762 n 7; *id.* at 773 (White dissenting). The objection to the licensing scheme thus could not have turned on the extent to which it curtailed expressive opportunities.

<sup>121</sup> The Court, in explaining why it recognizes facial challenges to standardless licensing schemes, often focuses on the way such schemes deter or “chill” certain ideas. As the Court said in *Lakewood*:

[The] licensor’s unfettered discretion, coupled with the power of prior restraint, intimidates parties into censoring their own speech, even if the discretion and power are never actually abused. . . . It is not difficult to visualize a newspaper . . . feeling significant pressure to endorse the incumbent mayor . . . to receive a favorable and speedy disposition on its permit application.

Hence, the claim might go, standardless licensing, in Meiklejohn's phrase, mutilates the community's thinking process. But I have reviewed this claim before and found it wanting, except to the extent that it reduces to a concern with improper purpose.<sup>122</sup> As a conceptual matter, content-based actions as well may improve as mutilate the community's thinking process. So too, then, with the standardless licensing schemes that facilitate such actions. If there is reason to think that as a practical (rather than a conceptual) matter standardless licensing more often will distort than improve public debate, that reason relates to the fear of illicit motive on the part of licensing officials.<sup>123</sup> The real question, then, concerns governmental motive; it is whether a rule against standardless licensing will identify and reduce the incidence of improperly motivated administrative decisions.

The rule against standardless licensing indeed serves this function of flushing out bad motive by establishing a safeguard against administrative action based on the content of expression. I have discussed already how and why impermissible motive more often taints content-based than content-neutral decisions.<sup>124</sup> This taint may affect administrative actions based on content no less than their legislative counterparts. Enforcing a prohibition on improperly motivated action thus would counsel adopting a presumption against any administrative decision based on the content of speech, which indeed the Court has done. But even more is needed to enforce the stricture against improper motive, for content-based administrative action (unlike content-based legislative action) is itself hard to identify. Suppose an official denies a license to a speaker under a statute specifying no standards; in the absence of an admission, a court cannot easily determine whether the official based her decision on the content of the speech (let alone whether she allowed impermissible motive to infect the decision).<sup>125</sup> To enforce the prohibition on con-

---

<sup>122</sup> See text accompanying notes 96-101.

<sup>123</sup> The Court's repeated invocation of chilling effects in this context, see note 121, thus derives from the fear of improper motive. Standardless licensing schemes chill certain ideas and not others, as each speaker considers what speech will advance and what speech will hinder her attempt to obtain a license. The reason to assume that the resulting disparate impact necessarily will have adverse effects on public discourse relates to the danger that officials will allow inappropriate factors—hostility, sympathy, or self-interest—to taint their decisions.

<sup>124</sup> See text accompanying notes 102-04.

<sup>125</sup> As the Court noted in *Saia*: "In this case a permit is denied because some persons were said to have found the sound annoying. In the next one a permit may be denied

tent-based action—which is itself a proxy for a finding of impermissible purpose—the Court needs yet a further prophylactic rule, commanding facial invalidation of all statutes providing for standardless licensing.

Such a rule decreases the danger of content-based licensing decisions, which in turn decreases the danger of improperly motivated licensing decisions, in two related ways.<sup>126</sup> First, by requiring that a licensing scheme contain specific, content-neutral standards, the rule directly promotes administrative decision making in accordance with those, and only those, criteria. Although licensing officials may take into account the content of speech even in the face of formal constraints, they are less likely to consider content if the authorizing statute spells out neutral criteria than if it is silent on the grounds for action. Second, the requirement of specific standards facilitates judicial review of the bases of licensing decisions. Although standards hardly ensure that courts will detect licensing decisions grounded in content, they at least provide, as the Court has noted, “guideposts” to aid the inquiry “whether the licensor is discriminating against disfavored speech”,<sup>127</sup> without such standards to serve as a measuring stick, “the difficulty of effectively detecting, reviewing, and correcting content-based censorship ‘as applied’” must increase.<sup>128</sup> Standards thus help to prevent and detect content distinctions and, even more critically, the impermissible motives that likely underlie them.

Simply put, the rule against standardless licensing statutes has emerged to prevent a legislature from displacing to administrators the power to make decisions regarding speech on the basis of a criterion—the content of expression—likely to involve impermissible motive.<sup>129</sup> The next rule I consider is similar, for

---

because some people find the ideas annoying. Annoyance at ideas can be cloaked in annoyance at sound.” 334 US at 562.

<sup>126</sup> See David A. Strauss, *The Ubiquity of Prophylactic Rules*, 55 U Chi L Rev 190, 196 (1988).

<sup>127</sup> *Lakewood*, 486 US at 758.

<sup>128</sup> *Id* at 759. At a later point, the Court recognized that the ultimate question was not whether the administrative action was based on content, but whether it was supported by an illegitimate motive. The Court stated that “without standards to bound the licensor, speakers denied a license will have no way of proving that the decision was unconstitutionally motivated, and, faced with that prospect, they will be pressured to conform their speech to the licensor’s unreviewable preference.” *Id* at 760. Note here how the issue of chill is related to the issue of motive. See notes 121, 123.

<sup>129</sup> These statutes also may contain a hint of illicit legislative motive. Legislators, after all, know the dangers of conferring unfettered discretion on administrators, and such discretion is not usually necessary to accomplish legitimate ends. Perhaps, then, a grant

it too demands strict scrutiny of a content-neutral law that effectively delegates authority to others to make content-based decisions, raising concerns of improper motive—although here the delegates are not so much administrative officials as members of the public.

## 2. Laws turning on communicative effect.

Some laws are content neutral on their face, but turn in operation on “communicative effect”—that is, on the way an audience responds to the content of expression.<sup>130</sup> A common example is a breach-of-the-peace statute, as applied to persons engaged in expressive activity.<sup>131</sup> In the typical case, a speaker expresses certain ideas, an audience makes known its displeasure, and police officers, fearful of public disturbance, arrest the speaker. The statute under which the arrest occurs makes no reference to the content of speech; it applies to whatever speech provokes, or tends to provoke, a hostile reaction. Yet the courts act as if the statute referred in express terms to the ideas that prompted the response, upholding the conviction only on a showing of necessity.<sup>132</sup> The courts effectively overlook the facial neutrality of the law and focus instead on what sparks its application—the reaction of an audience to a message.

At first glance, these cases seem mysterious under a motive-based theory, for breach-of-the-peace laws appear to raise few concerns of improper purpose on the part of the legislature. A desire to prevent violence counts as a legitimate, nonideological reason for restricting speech under the conception of improper motive used in this Article. Further, the risk seems small of hostility or self-interest creeping into the decision making process. If a statute were to prohibit specific ideas, on the ground that they provoked violence, we indeed would fear that dislike of those ideas influenced the legislature’s decision. But a breach-of-

---

of standardless discretion indicates a legislative desire for administrators to engage in ideological censorship, which itself counts as improper motive.

<sup>130</sup> For discussion of the notion of communicative effect, see Ely, *Democracy and Distrust* at 111-15 (cited in note 67). For discussion of the law’s treatment of content-neutral laws turning on communicative effect, on which my analysis is partly based, see Stone, 25 *Wm & Mary L. Rev.* at 234-39 (cited in note 15).

<sup>131</sup> See *Gregory v City of Chicago*, 394 US 111 (1969); *Cox v Louisiana*, 379 US 536 (1965). See also *Edwards v South Carolina*, 372 US 229 (1963) (common law breach of peace); *Cantwell v Connecticut*, 310 US 296 (1940) (same).

<sup>132</sup> See cases cited in note 131. In the hostile audience context, courts most often phrase the standard of review in terms of “clear and present danger,” but the standard differs not at all from that used to review content-based statutes.

the-peace statute, or other similar law, preselects no particular ideas as posing a danger, instead applying to whatever speech, in the actual event, threatens violence. (Indeed, such a law may apply not only to speech, but also to conduct posing a risk of disorder; such breadth usually decreases further the chance of illicit purpose.<sup>133</sup>) Thus, if the focus is on legislators' motives alone, the hostile audience cases seem wrongly decided.

This conclusion, however, might come too fast, even if the focus remains on legislative motives. Legislators, after all, may well know what ideas will provoke a hostile response; at the least, legislators can guess that the ideas most likely to do so will be those most generally seen as controversial or offensive. Similarly, legislators may well know what ideas potentially falling within the statute most often will capture the notice of law enforcement officials. This broad sense of how the law will operate may taint any legislator's consideration of the law, whether she shares or disputes the views of the public and officials. Of course, a legislator also may anticipate the effects of other content-neutral laws, but these effects usually will be so numerous, disparate, and crosscutting as to minimize their potential to corrupt her decision. The case may be different as to a law that operates only against ideas that raise the ire of the public.

An even greater concern involves the motives of law enforcement officials. Hostile audience laws raise, though to a lesser degree, the same broad problem as standardless licensing schemes: the capacity of officials, under such laws, to take action based on their views of ideas. These laws, of course, do not authorize action on this basis. Further, enforcement officials may import into the administration of any restriction on speech such ideological considerations. But a bit of history suggests that hostile audience laws are especially prone to this abuse, as in case after case, decade after decade, police officers have responded hastily, to say the least, to the risk of disorder caused by disfavored speech.<sup>134</sup> Nor is this history very surprising, given the vague standards contained in most breach-of-the-peace statutes, which make such laws more than usually subject to discriminatory enforcement.<sup>135</sup> The facility with which improper motive may

---

<sup>133</sup> See Section III.D.

<sup>134</sup> See *Gregory*, 394 US at 111 (civil rights protest); *Cox*, 379 US at 536 (same); *Edwards*, 372 US at 229 (same); *Feiner v New York*, 340 US 315 (1951) (speech endorsing racial equality and criticizing public officials); *Chaplinsky v New Hampshire*, 315 US 568 (1942) (speech attacking religion and government); *Cohen v California*, 403 US 15 (1971) (speech attacking the draft).

<sup>135</sup> For example, the statute in *Cohen* prohibited any person from "maliciously and

taint—and the frequency with which it has tainted—the enforcement of hostile audience statutes thus provides another reason for treating such statutes with suspicion.

Most important, a hostile audience law, as applied, may constitute an improper delegation of authority to the public to suppress messages it disfavors. Consider how such a law operates in practice. A public disturbance of some kind triggers the restriction, but the content of the speech usually triggers the disturbance, and ideological considerations usually enter into the public's response to the speech's content. The constitutional difficulty grows out of this nexus between the government's action and the majority's viewpoint. First Amendment doctrine treats identically regulation stemming from incumbent officials' hostility toward ideas and regulation stemming from (or deferring to) the analogous attitudes of the majority.<sup>136</sup> As the government cannot itself censor citizens for ideological reasons, so too it cannot authorize one group of citizens to censor another.<sup>137</sup> This principle leaves the government free to restrict speech provoking a hostile response when necessary to prevent violence, but bars the government from restricting such speech in reflection of or deference to majoritarian bias. Stringent scrutiny of each application of a hostile audience law separates the one kind of action from the other; it ensures that the government has not simply empowered the public to restrict speech on the basis of content, with all the likelihood of ideological censorship such a restriction implies.<sup>138</sup>

The key to the analysis of this Section is first, the functional equivalence between statutes referring to content and statutes turning on communicative impact and second, the relation between content discrimination and impermissible motive. The first point is that laws turning on communicative impact (much like laws establishing standardless licensing schemes), though con-

---

willfully disturbing the peace or quiet of any neighborhood or person . . . by . . . offensive conduct." 403 US at 16. Similarly, the statute in *Gregory* made it unlawful for any person to make "any improper noise, riot, disturbance, breach of the peace, or diversion tending to a breach of the peace." 394 US at 116 (Black concurring). See note 117 on the relationship between vagueness and improper motive.

<sup>136</sup> See text accompanying notes 49-51.

<sup>137</sup> See Geoffrey R. Stone, *Reflections on the First Amendment: The Evolution of the American Jurisprudence of Free Expression*, 131 Proc Am Phil Soc'y 251, 258 (1987).

<sup>138</sup> The fighting words doctrine constitutes an exception to the principle that statutes turning on communicative impact, though facially content neutral, receive heightened scrutiny because they delegate to the public the ability to restrict speech based on content. I consider in Section III.C.1 why fighting words laws receive minimal scrutiny.

tent neutral on their face, allow content-based actions in application, whether by enforcement officials or the public. The second point is that content-based action—whether on the face of a law or as it applies—raises the fear of improper motive. I now turn to one last case in which that fear comes into play, notwithstanding a statute's facial neutrality.

### 3. Laws "equalizing" the speech market.

In what has become one of the most castigated passages in modern First Amendment case law, the Court pronounced in *Buckley v Valeo* that "the concept that government may restrict the speech of some elements of our society in order to enhance the relative voice of others is wholly foreign to the First Amendment . . ."<sup>139</sup> The Court made this statement in the course of invalidating expenditure ceilings in a campaign finance law, one justification for which was that they "equaliz[ed] the relative ability of individuals and groups to influence the outcome of elections."<sup>140</sup> But the statement has ramifications far beyond the area of campaign finance. It applies as well to a wide variety of schemes designed to promote balance or diversity of opinion. So, for example, the statement could have explained the Court's earlier holding in *Miami Herald Publishing Co. v Tornillo* disapproving a "right of reply" statute that required a newspaper to grant space in its pages to any political candidate whose record the newspaper had criticized.<sup>141</sup> So too the statement could summarize the view of the four dissenting Justices in *Turner Broadcasting System, Inc. v FCC* that "must-carry" rules, requiring cable operators to provide access to local broadcast television stations, violate the First Amendment.<sup>142</sup>

---

<sup>139</sup> 424 US 1, 48-49 (1976). Not surprisingly, scholars who advocate a model of the First Amendment focusing on the quality of public debate have attacked the Court's statement most harshly. See Sunstein, *Democracy and the Problem of Free Speech* at 94-101 (cited in note 19); Fiss, 71 Iowa L Rev at 1423-25 (cited in note 97).

<sup>140</sup> *Buckley*, 424 US at 48. The Court also has refused to accept this justification (often quoting *Buckley*) in other cases involving campaign finance regulations. See, for example, *First National Bank v Bellotti*, 435 US 765, 790-91 (1978) (invalidating a statute restricting corporate contributions and expenditures in referendum campaigns); *Citizens Against Rent Control v City of Berkeley*, 454 US 290, 295-96 (1981) (invalidating an ordinance limiting contributions to committees taking positions on referenda).

<sup>141</sup> 418 US 241 (1974).

<sup>142</sup> 114 S Ct 2445, 2475-81 (1994) (O'Connor concurring in part and dissenting in part). It is possible to argue that the statement in *Buckley* applies to an expenditure limit, but not to a right-of-reply statute or a must-carry rule, because the former is a direct restriction on speech, whereas the latter are merely rules regulating access to expressive

The Court's commitment to the principle of *Buckley* has lapsed on some occasions. After the Court invalidated *Buckley's* expenditure ceilings (which applied to individuals and groups alike), it accepted as an adequate justification for a statute limiting a corporation's political expenditures that corporate wealth could cause "distortion" and "unfairly influence elections."<sup>143</sup> And a few years before the Court struck down *Tornillo's* right-of-reply statute, the Court approved the FCC's fairness doctrine, which operated in a similar manner and arose from identical concerns relating to diversity and balance.<sup>144</sup> Finally, a slim majority in *Turner* upheld the must-carry rules imposed on cable operators—although, notably, by refusing to view the rules as an attempt to achieve an appropriate mix of ideas and information.<sup>145</sup> Still, *Buckley's* antiredistribution principle has con-

---

channels. But this proposed distinction collapses on reflection. The so-called access regulations at issue in *Tornillo* and *Turner* prevent the speaker from using the space transferred to others for her own preferred expression and thus may be said to constitute a direct restriction. Conversely, the so-called direct restriction in *Buckley* operates to provide others with greater access to the sphere of campaign expression, while allowing the speaker some continuing ability to express herself, and thus may be said to constitute an access regulation. I do not wish to say that all access rules are like direct restrictions and vice versa. But wherever a given resource is limited, it does no good to try to draw too sharp a line between them; providing access will restrict speech and restricting speech will provide access. The rationale in *Buckley* applies to either.

<sup>143</sup> *Austin v Michigan Chamber of Commerce*, 494 US 652, 660 (1990). The Court tried to distinguish *Austin* from *Buckley*, principally on the ground that corporate wealth derives from privileges bestowed on corporations by the government. *Id.* But this argument fails, because individual wealth also derives from governmental action. What the Court recognized in *Austin* is only what is true in every case: direct regulation of speech occurs against a backdrop of law that, while not referring to speech, goes far toward structuring the sphere of public expression. The question in every case is whether the government may use direct regulation of speech to redress prior imbalances.

<sup>144</sup> *Red Lion Broadcasting Co. v FCC*, 395 US 367 (1969). The Court in *Tornillo* never mentioned *Red Lion*, but presumably believed the case did not apply because it involved broadcasting, rather than the print media. The *Red Lion* Court had reasoned that different First Amendment standards govern broadcasting because of the (then existing) scarcity of radio frequencies. 395 US at 386-88.

<sup>145</sup> The Court interpreted the must-carry rules as an attempt "not to favor programming of a particular subject matter, viewpoint, or format, but rather to preserve access to free television programming for the 40 percent of Americans without cable." *Turner*, 114 S Ct at 2461. Congress's concern, according to the Court, was that cable operators, in the absence of must-carry provisions, would refuse carriage to local broadcast stations and that this refusal would destroy the viability of free broadcast television. The dissent objected to this reading of the must-carry rules, arguing that they instead were an attempt to promote diversity of information on the airwaves, by favoring speakers who would provide significant amounts of educational, local news, and public affairs programming. *Id.* at 2476-77 (O'Connor concurring in part and dissenting in part). What matters here is that the Court understood the must-carry rules in such a way as to obviate any conflict between those rules and the *Buckley* principle. If the rules were not designed to improve the content of the speech market, then they could not conflict with a prohibition

tinuing importance: no Justice on the current Court would dispute the claim—even if some Justices would dispute applications of it—that the government may not restrict the speech of some to enhance the speech of others.<sup>146</sup> The question that remains concerns the basis of this principle: what view of the First Amendment accounts for the Court's refusal to allow, by means of restrictions, the redistribution of expression?

The audience-based conception of the First Amendment does not; indeed, that conception calls for the converse of the *Buckley* principle. If what is essential, to recall Meiklejohn's phrase, is that "everything worth saying shall be said,"<sup>147</sup> then the First Amendment often would permit—indeed require—the reallocation of speech opportunities. The realm of public expression may have too much of some kinds of speech, too little of others; some speakers may drown out or dominate their opposite numbers. Self-conscious redistribution of expressive opportunities seems the most direct way of correcting these defects and achieving the appropriate range and balance of viewpoint.

Neither does the speaker-based conception of the First Amendment explain the *Buckley* principle.<sup>148</sup> The speaker-based model counsels in every case a weighing of the scope of the restriction on speech against the importance of the asserted state interest. Under this analysis, the Court in *Buckley* would have asked whether the interest in promoting diversity of opinion outweighed the loss of expression resulting from expenditure ceilings. But the Court in *Buckley* followed no such analysis when it considered the government's redistributive interest.<sup>149</sup> The

---

on the government's undertaking such a project.

<sup>146</sup> Although the matter is not free from doubt, the antiredistribution principle of *Buckley* may well be subject to a "necessity" defense, allowing the Court to uphold redistributive speech policies upon a showing of need similar to that required for other suspect regulations of speech, such as content-based restrictions. The Court suggested as much in *Bellotti* when it said that if arguments relating to the "undue influence" of corporate speech "were supported by record or legislative findings that corporate advocacy threatened imminently to undermine democratic processes, thereby denigrating rather than serving First Amendment interests, these arguments would merit our consideration." 435 US at 789. Perhaps, too, the Court's decision in *Red Lion* emerged from the view that the fairness doctrine satisfied a test of necessity given the then existing scarcity of broadcast frequencies. See Geoffrey R. Stone, *Autonomy and Distrust*, 64 U Colo L Rev 1171, 1178 (1993). In any event, regulations thought suspect on the ground of their redistributive justification probably face a rebuttable, rather than an irrebuttable, presumption against them.

<sup>147</sup> Meiklejohn, *Political Freedom* at 26 (cited in note 15).

<sup>148</sup> Owen Fiss has suggested that the *Buckley* principle emerges from the speaker-based model, which he calls the autonomy principle. See Fiss, 71 Iowa L Rev at 1423 (cited in note 97). I disagree with this view of the *Buckley* principle's underpinnings.

<sup>149</sup> The Court did use this analysis in evaluating the distinct governmental interest in

Court said not that this interest was insufficient, but that its very assertion conflicted with fundamental premises of the First Amendment. No matter how little the ceilings limited expression or how much they promoted diversity of opinion, the Court could not countenance their existence. Such an analysis fits poorly with the speaker-based model because it rests not on the extent of, but on the justification for, the restriction.

The motive-based model of the First Amendment, last left standing though it is, appears at first to explain the *Buckley* principle no better than the others. It is true that *Buckley* is all about reasons; the decision bars the government from asserting or relying on an interest in redistribution to support a limit on expression. But the interest that *Buckley* declares improper is not congruent with—indeed, is almost the converse of—the concept of illicit motive I have developed. The justification barred in *Buckley*, far from relating to the decision maker's own ideology or self-interest, relates to her desire to construct a speech market in which citizens can choose, in an informed way, among many competing ideas, benefiting many different interests. How, then, does the *Buckley* principle fit with the concept of improper motive prevalent in the rest of First Amendment doctrine?

The answer to this question involves viewing the *Buckley* principle as an evidentiary tool designed to aid in the search for improper motive, much like the presumption against content-based restrictions. The *Buckley* principle emerges not from the view that redistribution of speech opportunities is itself an illegitimate end, but from the view that governmental actions justified as redistributive devices often (though not always) stem partly from hostility or sympathy toward ideas—or, even more commonly, from self-interest.<sup>150</sup> The *Buckley* principle thus serves a function of proof—once again, to uncover, in the absence of a

---

preventing corruption. There, the Court held, as an outgrowth of balancing, that the interest was "inadequate to justify" the restriction. *Buckley*, 424 US at 45. That approach is consistent with a speaker-based model. It also may represent an attempt, as with review of any content-neutral law, to discover the slight chance of impermissible motive.

<sup>150</sup> This interpretation of the *Buckley* principle explains why the Court—as suggested by the holding in *Red Lion*, 395 US at 367, and dictum in *Bellotti*, 435 US at 789—will accept a redistributive justification on a showing of necessity. See note 146. The ability to rebut the presumption of unconstitutionality flowing from a redistributive justification by showing that redistribution is desperately needed indicates that pursuing the goal of diversity is not itself illegitimate; the presumption of unconstitutionality must attach to the redistributive justification because that justification tends to be linked (in a way that a showing of need will disprove) to other, impermissible ends of government.

direct method for doing so and in a necessarily inexact manner, improperly motivated governmental actions.

When laws justified as redistributive devices draw content-based lines, the evidentiary function of the *Buckley* principle becomes clear, though also superfluous. Assume Congress enacts a law limiting speech in support of the Republican Party, on the ground that speech of this kind drowns out speech promoting other political views. Or assume Congress, to take a serious proposal, passes a law limiting sexually explicit speech promoting the subordination of women, on the ground that such speech dominates existing avenues of communication and, indeed, silences the speech of those opposed to it. In both cases, the Court might invoke the *Buckley* principle; and in both cases that principle would serve the same motive-hunting function as the (also applicable) presumption against content-based action. The justifications offered for these laws might be valid. But such laws, like all other content-based restrictions, pose a significant danger of arising from the different desire on the part of officials to insulate themselves or their ideas from challenge. In these cases, to state the slogan of *Buckley* is only to say that when officials restrict a particular idea—with the ostensible goal of increasing diversity, as with any other proper aim—the Court will not trust them to fend off improper considerations of hostility, sympathy, or self-interest.

The question remains, however, why the Court should treat as especially suspicious content-neutral regulations of speech—such as the regulations in *Buckley*—that are justified in terms of achieving diversity.<sup>151</sup> If the *Buckley* principle serves

---

<sup>151</sup> It is often difficult to classify schemes justified in terms of diversity as content based or content neutral. *Buckley* provides an example. I treat the expenditure ceilings in *Buckley* as content neutral, as is the norm, because they cover spending for expression supporting all political candidates. See Tribe, *American Constitutional Law* § 13-27 at 1132-36 (cited in note 67). It could be argued, however, that the ceilings imposed a subject-matter limitation because they applied only to spending in political campaigns. A similar question arises with respect to the right-of-reply laws in *Red Lion* and *Tornillo*, which also applied equally to all sides of a subject, but only to subjects of a certain kind. And cases involving diversity schemes often raise the problem of how to categorize so-called speaker-based restrictions on expression. The must-carry rules in *Turner*—like the corporate speech restrictions in *Bellotti* and *Austin*—primarily distinguished on the basis of the identity of the speaker, rather than the content of speech. Particular speakers, though, tend to say particular things; local broadcast stations, for example, usually carry programming on local issues. These facially speaker-based provisions thus raise the question—inadequately addressed in *Turner*—whether and when the Court should associate a given speaker with a given idea, thus treating the regulation as content based. See generally Stone, 25 Wm & Mary L Rev at 244-51 (cited in note 15). Of course, if the Court

as a means for flushing out illicit motive, then the answer must relate to some special characteristic(s) of these regulations that affect the motive inquiry. And in fact, the nature of these regulations, as compared with other content-neutral regulations, creates two problems (similar to those posed by standardless licensing schemes): first, that governmental officials (here, legislators) more often will take account of improper factors, and second, that courts will have greater difficulty detecting the presence of such tainted deliberations.

The increased probability of taint arises, most fundamentally, from the very design of laws directed at equalizing the realm of public expression. Unlike most content-neutral regulations, these laws not only have, but are supposed to have, content-based effects; their *raison d'être* is to alter the mix of ideas—or, at least, of speakers, who tend to be associated with ideas—in the speech market. Given this function, these laws will have not the diverse, diffuse, and crosscutting content-based effects usually associated with content-neutral laws, but a set of targeted and coherent effects on ideas and speakers. This set of focused effects renders a law directed at equalization nearly as likely as a facially content-based law, and much more likely than most facially content-neutral laws, to stem from improper motive. In considering such a law, a legislator's own views of the ideas (or speakers) that the equalization effort means to suppress or promote may well intrude, consciously or not, on her decision-making process. The law thus raises grounds for suspicion.

The likelihood of impermissible taint increases further because of the nature of the harm claimed to justify the action. This harm—a disproportionate access or undue influence—has an amorphous aspect to it that often (though not always) will make it peculiarly difficult to measure.<sup>152</sup> This does not mean that legislators cannot debate rationally the size of the harm and cannot reach what appears to be—and in fact might be—a reasonable decision to counteract it. But because the harm has a fuzzy quality, the decision whether the harm justifies a limit on speech becomes especially susceptible to the infiltration of illicit factors. It is the rare person who can determine whether there is

---

treats identically all restrictions of speech justified in terms of diversity, regardless whether they are content based or content neutral, the Court will never have to confront these labeling issues. The question on the table is whether there is any other reason for providing the same treatment to all such restrictions.

<sup>152</sup> See text accompanying note 100.

"too much" of some speech (or speakers), "too little" of other speech (or speakers), without any regard to whether she agrees or disagrees with—or whether her own position is helped or hurt by—the speech (or speakers) in question.

Indeed, all the laws directed at equalization that the Court has considered, whether classified as facially content based or content neutral, raise questions as to the motives of the enacting legislatures. Campaign finance laws like those in *Buckley* easily can serve as incumbent-protection devices, insulating current officeholders from challenge and criticism.<sup>153</sup> When such laws apply only to certain speakers or subjects, the danger of illicit motive becomes even greater; for example, the law in *First National Bank v Bellotti*,<sup>154</sup> which barred corporations from spending money in referendum campaigns, almost surely arose from the historic role of corporate expenditures in defeating referenda on taxation.<sup>155</sup> Similarly, a right-of-reply law like the one in *Tornillo*—applicable only to political candidates, albeit to all of them—may have stemmed not from the desire of officials to enhance the quality of public debate, but from their wish to get the last word whenever criticized.<sup>156</sup> If this law did not quite prohibit seditious libel, it came close. And the must-carry rules in *Turner* may have emerged more from the yen of politicians for local publicity than from their wish either to preserve free television or to enhance public discourse.<sup>157</sup> All these examples suggest the same point: there may be good reason to distrust the

---

<sup>153</sup> See, for example, Lillian R. BeVier, *Money and Politics: A Perspective on the First Amendment and Campaign Finance Reform*, 73 Cal L Rev 1045, 1080 (1985); *Austin*, 494 US at 692-93 (Scalia dissenting) (suggesting that "with evenly balanced speech incumbent officeholders generally win").

<sup>154</sup> 435 US 765 (1978).

<sup>155</sup> See *id.* at 826-27 n 6 (Rehnquist dissenting). Even in voting to uphold the law, then Associate Justice Rehnquist admitted: "[A] very persuasive argument could be made that the General Court, desiring to impose a personal income tax but more than once defeated in that desire by . . . corporate expenditures in opposition to such a tax, simply decided to muzzle corporations on this sort of issue so that it could succeed in its desire." *Id.* See also *Austin*, 494 US at 692 (Scalia dissenting) (noting that a limit on corporate expenditures in support of political candidates might have been "trying to give unincorporated unions . . . political advantage over major employers").

<sup>156</sup> The fairness doctrine upheld in *Red Lion* included a similar provision allowing candidates a right to reply to any editorial endorsing an opponent. This provision, however, formed only part of a more general policy allowing any person or group to respond to personal attacks and requiring coverage of all sides of public issues. See *Red Lion*, 395 US at 373-75. The greater generality of this law, as compared with the one in *Tornillo*, suggests a lesser reason to distrust the legislative process.

<sup>157</sup> Local broadcast stations, which are the stations cable operators must carry, usually provide greater coverage of local politicians than other programmers do.

motives of politicians when they apply themselves to reconstructing the realm of expression.

The problem is compounded by the difficulty of detecting, through use of the standard of review applied to content-neutral laws, the presence of improper taint in statutes ostensibly aimed at equalizing expression.<sup>158</sup> The unusual impotence of the standard in this context relates to the absence of any clear criteria for deciding what state of public debate constitutes the ideal and how far current discourse diverges from it. If a court cannot make these fine judgments, then it also cannot conclude that a law supposedly enacted to ease distortion in fact fails to advance this interest or to do so in a narrowly tailored manner.<sup>159</sup> In almost any case, the government can defend a statute on the ground that it alleviates distortion; and in almost no such case—even when the asserted purpose is a pretext—will the standard of review applied to content-neutral laws command the statute's invalidation. The usual motive-hunting mechanism fails to operate.

The ease with which improper purpose can taint a law directed at equalizing expression together with the difficulty a reviewing court will have in detecting this taint account for the Court's approach in cases like *Buckley*. The reason for the approach is not that the goal of equalization itself conflicts with the First Amendment (though the Court often speaks in this manner); to the contrary, when Justice Scalia labeled the objective of equalization "unqualifiedly noble,"<sup>160</sup> he may have used the term tongue-in-cheek, but he also used it aptly. The reason for the approach is instead that the goal of equalization often and well conceals what does conflict with the First Amendment: the passage of laws tainted with ideological, and especially with self-interested, motivations. Again in Justice Scalia's words: "The incumbent politician who says he welcomes full and fair debate is

---

<sup>158</sup> As noted earlier, this standard usually works to flush out content-neutral laws tainted by illicit purpose. See text accompanying notes 108-09.

<sup>159</sup> See BeVier, 73 Cal L. Rev at 1090 (cited in note 153). Yet another problem in testing the fit of a statute of this kind arises from the fact that the governmental interest at stake relates exclusively to speech. As I discuss in later Sections, when legislators can advance an interest both by regulations restricting speech and by regulations restricting nonexpressive conduct, courts can explore the bona fides of a limitation on speech by considering whether and how the government has tried to promote the interest in non-speech-related ways. But because the interest in equalizing the speech market is about nothing other than speech, this technique is unavailable to courts in the cases I have been discussing, thus further impeding the search for impermissible motive.

<sup>160</sup> *Austin*, 494 US at 692 (Scalia dissenting).

no more to be believed than the entrenched monopolist who says he welcomes full and fair competition."<sup>161</sup> The harsh treatment of laws directed at correcting distortion, even when these laws are framed in content-neutral language, arises from the fear that if the usual standards of review applied, legislators would use these laws as a vehicle for improper motive, and courts would bless what the First Amendment proscribes.

Laws directed at equalizing speech thus join the list of laws that, although facially content neutral, demand strict scrutiny because of heightened concerns relating to improper purpose. In all these cases, the laws in some way rest on or allow decisions based on content, though no reference to content appears on their face. In so doing, the laws raise the same suspicions of improper purpose as do facially content-based laws, and for much the same reasons. The Court thus treats these laws in a strict manner—presuming improper taint, though giving the government a chance to rebut this presumption. The doctrine, again, responds to the desire to discover impermissible purpose—and to the necessity of relying on overinclusive and underinclusive categorical rules to accomplish this object.

### C. "Safe" Content-Based Laws

If some facially content-neutral laws pose a serious risk of stemming from improper purpose, perhaps the converse is also true: that some facially content-based laws pose only a slight danger of taint, so that weakening the presumption against them becomes appropriate. In this Section, I look at First Amendment law's two great exceptions to the rule against content-based restraints—the recognition of low-value categories of speech and the doctrine of secondary effects—and ask whether a motive-based approach can explain them. I conclude that this approach, although not answering all questions in these two contexts, provides the most coherent general account of prevailing doctrine.

#### 1. Low-value speech.

First Amendment law is replete with content distinctions that do not count as content distinctions because they disfavor speech found by the Court to have little (or no) constitutional value and thus to receive little (or no) constitutional protection.

---

<sup>161</sup> *Id.*

The government may ban obscenity or child pornography wholesale without implicating First Amendment principles.<sup>162</sup> So too for fighting words, as the *R.A.V.* Court reiterated.<sup>163</sup> Less simply, the regulation of commercial speech receives minimal constitutional scrutiny if the speech is false or misleading, intermediate scrutiny in all other cases.<sup>164</sup> Least simply, libel law is subject to a bewildering variety of constitutional standards, which (roughly speaking) become less exacting as the plaintiff and the subject matter become less "public."<sup>165</sup> All of these disfavored categories are based on the content of speech; some, at least arguably, are based on its viewpoint.<sup>166</sup>

These "low-value" categories have a flipside—a category of converse content, though of a more informal nature, which may place them in valuable perspective. The Court sometimes has stated that the First Amendment protects most strongly speech

---

<sup>162</sup> See *Miller v California*, 413 US 15, 23 (1973) (obscenity); *New York v Ferber*, 458 US 747, 758 (1982) (child pornography). In addition, the Court has indicated that other sexually explicit speech is entitled to less than a full measure of constitutional protection, although the precise standard is uncertain. See *FCC v Pacifica Foundation*, 438 US 726, 744-48 (1978); *Young v American Mini-Theatres, Inc.*, 427 US 50, 62-63 (1976); *City of Renton v Playtime Theatres, Inc.*, 475 US 41, 48-50 (1986).

<sup>163</sup> The fighting words doctrine originated in *Chaplinsky v New Hampshire*, 315 US 568 (1942); although in recent years the doctrine's continued vitality seemed in doubt, the Court in *R.A.V.*, 505 US at 382-83, treated it as fully settled law. From the beginning, the Court defined the category of fighting words not by reference to specific content, but by reference to the typical addressee's response; as argued previously, however, this difference should have no operational significance, given that an audience's response to speech itself turns on content. See text accompanying notes 136-38.

<sup>164</sup> See *Central Hudson Gas & Elec. Corp v Public Serv Comm'n*, 447 US 557, 562-64 (1980).

<sup>165</sup> See *Gertz v Robert Welch, Inc.*, 418 US 323, 342-43 (1974); *Dun & Bradstreet, Inc. v Greenmoss Builders, Inc.*, 472 US 749, 760-61 (1985) (plurality opinion). Many complexities to one side, public officials and figures must prove actual malice on the part of defendants to recover damages; private figures complaining about a statement of public concern must show negligence to recover compensatory damages and actual malice to recover punitive and presumed damages; and private figures complaining about a statement of purely private concern need show only what state law requires for punitive and presumed damages and probably for compensatory damages as well.

<sup>166</sup> Catharine MacKinnon has argued, for example, that the categories of obscenity and child pornography reflect a kind of viewpoint discrimination because the speech falling within these categories likely expresses a single (disfavored) viewpoint about sexual matters. See Catharine A. MacKinnon, *Feminism Unmodified: Discourses on Life and Law* 212 (Harvard 1987). Cass Sunstein has strengthened the argument by noting that the test of obscenity invokes community standards of offensiveness. See Cass R. Sunstein, *Neutrality in Constitutional Law (With Special Reference to Pornography, Abortion, and Surrogacy)*, 92 Colum L Rev 1, 28-29 (1992). A similar argument of viewpoint bias might apply to the fighting words category, given that it is often the viewpoint (and not merely the generalized "content") of speech that will prompt an average addressee to fight.

on political matters; it is in this area of expression, in the words of the Court, that the First Amendment achieves its "fullest and most urgent application."<sup>167</sup> Statements of this kind, it might be argued, have had no real import. As a rule, no separate standard reigns for speech on political affairs—or even, to dilute the category a bit, speech on matters of public interest.<sup>168</sup> The Court's pronouncements thus may have served only a packaging function: to provide added (but unnecessary) support for decisions that would have been reached regardless. But there is reason to think, contrary to this interpretation, that the political character of expression in fact affects the Court's decisions, even if in a more subtle fashion than by raising the formal constitutional standard. In recent decades, the Court almost never has upheld a regulation of political speech. Perhaps more tellingly, almost all of the landmarks of First Amendment law—the classic cases that set the tone and provide the focus for analysis of free speech questions—arise out of governmental attempts to restrict speech of an obviously political nature.<sup>169</sup>

---

<sup>167</sup> *Monitor Patriot Co. v Roy*, 401 US 265, 272 (1971); *Minneapolis Star & Tribune Co. v Minnesota Comm'r of Revenue*, 460 US 575, 585 (1983). See also *Bellotti*, 435 US at 776-77; *Buckley*, 424 US at 14-15; *Mills v Alabama*, 384 US 214, 218 (1966); John Paul Stevens, *The Freedom of Speech*, 102 Yale L J 1293, 1306-07 (1993).

<sup>168</sup> The law on the speech of public employees provides one exception to this principle. The Court has ruled that the First Amendment prohibits discharging an employee because of her speech only when the speech relates to "matters of public concern." See *Connick v Myers*, 461 US 138, 142 (1983). The Court's libel jurisprudence also has incorporated, in an on-again-off-again way, a separate standard for speech on matters of public interest. At one time, all (but only) statements of "general or public interest" received the highest level of constitutional protection from the prosecution of libel actions. *Rosenbloom v Metromedia, Inc.*, 403 US 29, 48-49 & n 17 (1971) (plurality opinion). The Court in *Gertz* rejected this approach in favor of a method that makes the level of protection dependent on the nature of the plaintiff rather than the subject matter of the speech. The question whether the plaintiff is a public or a private figure, however, partly depends on whether the allegedly libelous speech stemmed from her participation in a "public controvers[y]"—a decision scarcely different from the decision whether the allegedly libelous speech concerned a public issue. See *Gertz*, 418 US at 343-46. Moreover, in a later case, the Court reintroduced such an inquiry by dividing libel suits brought by private figures between those involving matters of "purely private concern" and those involving other subjects. See *Dun & Bradstreet*, 472 US at 758-60 (plurality opinion). See generally Cynthia L. Estlund, *Speech on Matters of Public Concern: The Perils of an Emerging First Amendment Category*, 59 Geo Wash L Rev 1 (1990).

<sup>169</sup> The list is a long one, including the great dissenting opinions of Holmes in *Abrams v United States*, 250 US 616, 624 (1919), and *Gillow v New York*, 268 US 652, 672 (1925), the equally great concurring opinion of Brandeis in *Whitney v California*, 274 US 357, 372 (1927), and the majority opinions in *Near v Minnesota*, 283 US 697 (1931), *New York Times Co. v Sullivan*, 376 US 254 (1964), *Brandenburg v Ohio*, 395 US 444 (1969), *Cohen v California*, 403 US 15 (1971), the Pentagon Papers case, otherwise known as *New York Times Co. v United States*, 403 US 713 (1971), and the flag burning cases of *Texas v Johnson*, 491 US 397 (1989), and *United States v Eichman*, 496 US 310 (1990). For a

This schema—the formal delineation of low-value categories on the one hand, the informal elevation of political speech on the other—plausibly derives from more than one model of the First Amendment. Although the speaker-based conception fits awkwardly with the doctrine, both the audience-based model and the government- or motive-based model substantially, though not entirely, support it. Thus, this aspect of First Amendment law, unlike others I have noted, is not solely explicable as a tool for discovering improper motive. But as I explain, the audience-based account of low-value categories creates conflict—or at least tension—between the recognition of those categories and other fundamental aspects of First Amendment law. In contrast, the motive-based account of these categories renders them harmonious with the rest of the doctrine.

A speaker-based approach could explain low-value categories only if speech of the disfavored kinds confers less of value on a speaker than does speech receiving full protection. Low-value speech must promote autonomy or self-realization or self-development—the values, variously conceived and phrased, that a speaker gains from communicating—less well than does other speech. Further, political speech, to explain the flipside of the doctrine, must promote these values better.

This account of the doctrine seems peculiar. It is, after all, not clear what kind of speech does the greatest good for speakers, or best promotes their interests; it is not even clear what criteria we would use to think about this question.<sup>170</sup> One scholar has called the speech in *Chaplinsky v New Hampshire*<sup>171</sup> “a personal catharsis, . . . a means to vent his frustration[,] . . . a significant means of self-realization.”<sup>172</sup> Another proponent of the speaker-based model has written that “materials that most people view as pornographic may play an important role in some people’s self-fulfillment and self-expression.”<sup>173</sup> Given the uncer-

---

popular history of First Amendment doctrine that focuses on these cases, weaving them into a coherent First Amendment tradition, see Anthony Lewis, *Make No Law: The Sullivan Case and the First Amendment* (Random House 1991).

<sup>170</sup> See Sunstein, *Democracy and the Problem of Free Speech* at 141 (cited in note 19) (“If we protect speech because people want to talk, it is not easy to come up with standards by which to distinguish among different kinds of talk.”).

<sup>171</sup> 315 US 568 (1942).

<sup>172</sup> Martin H. Redish, *Freedom of Expression: A Critical Analysis* 56 (Michie 1984).

<sup>173</sup> C. Edwin Baker, *Human Liberty and Freedom of Speech* 69 (Oxford 1989). The same scholar has argued that commercial speech ought not to receive constitutional protection because such speech “reflects market forces” rather than manifests individual choice. *Id.* at 196. These are the kinds of distinctions that would have to be made in

tainty surrounding these issues, the most appropriate course would place in the speaker's own hands the question what kind of speech has value to her, by freeing her to choose among expressive activities.<sup>174</sup> Indeed, to the extent governmental officials influence this choice, they may stifle the very aspect of expression—the speaker's decision as to what kind of speech has greatest value to her—most closely related to self-realization. But even placing this claim to one side, it seems a dubious project to expropriate, on grounds only of the speaker's own interests, the speaker's decision, expressed through the act of speech itself, as to what forms of speech have value to her.

An audience-based perspective does better in accounting for First Amendment doctrine's categorization of speech according to value. If the goal of a free speech system is to provide individuals (especially in their roles as citizens) with the range of opinion and information that will enable them to arrive at truth and make wise decisions, then a tiered system of speech, of the kind the Court has created, seems appropriate. Some speech does not enrich (may even impoverish) the sphere of public discourse. Other speech contributes to reasoned deliberation on matters of public import. Under the audience-based approach, it would be perverse to treat these disparate forms of speech identically. Thus emerges a multitiered system.

And thus might emerge *this* multitiered system, which represents a plausible statement of the contribution of different kinds of speech toward creating a healthy sphere of discourse. The extreme respect shown to political speech comports with the strand of the audience-based model that highlights the need to give individuals the information required to fulfill the role of sovereign citizen. The treatment of libel rests on the premise that false statements of fact have no value because they prevent listeners from gaining accurate information; the varying standards applied to different kinds of libel suits then reflect roughly the value of the true speech that these suits chill.<sup>175</sup> The treatment

---

developing a tiered First Amendment through a speaker-based model.

<sup>174</sup> This does not mean that courts may not restrict the speech so chosen; they may do so if the speech poses the requisite danger of harm. But under a speaker-based approach, for the reasons stated in the text, courts usually should not assign varying values to varying forms of expression, so that disparate showings of harm are required.

<sup>175</sup> The correlation here is admittedly crude. If libel standards directly reflected the value of speech to public discourse, they would vary with the subject matter of the speech, rather than with the character of the plaintiff. One scholar who favors an audience-based approach proposes this reform of the law of defamation. See Sunstein, *Democracy and the Problem of Free Speech* at 159-62 (cited in note 19). As noted previously, however, the law

of false and misleading commercial speech arises from a similar sense that speech that operates through deception cannot serve audience interests. And the exclusion of obscenity and fighting words reflects a judgment that these forms of speech are themselves not "reasoned" and thus cannot aid reasoned discourse.

Some aspects of the current multitiered system, however, fit oddly with the audience-based model. The near absolute protection given to false but nondefamatory statements of fact outside the commercial realm is peculiar from this perspective; even a concern with chilling true speech would not explain such sweeping protection of speech that disserves understanding. Conversely, the reduced protection granted to truthful commercial speech, when compared with speech not only on politics but on other subjects, raises questions; to justify this treatment, the audience-based model needs a plausible theory of why learning about commercial matters and making sound commercial decisions has only insubstantial value to the public. Even the noncoverage of fighting words and obscenity is less than certain under this model. Perhaps personal invective adds something to public dialogue, precisely because of its earthy quality. Or perhaps some materials labeled obscene enrich understanding of sexual matters.

Even to the extent that the audience-based model makes sense of low-value categories, it has trouble bringing this doctrinal feature into line with the rest of First Amendment law. Many scholars have objected on this score to the creation and use of low-value categories. In the words of one, commenting on the reasoning of *Chaplinsky*, the inquiry whether certain kinds of speech are of "slight social value as a step to truth" compels the Court to make "value judgments concerned with the content of expression, a role foreclosed to it by the basic theory of the First Amendment."<sup>176</sup> Or, in the words of another, "the very concept of low-value speech is an embarrassment to first amendment orthodoxy. To say that government cannot suppress speech unless the speech is of low value sounds like a parody of free speech theory."<sup>177</sup> "Embarrassment" or "parody" may be too strong, but

---

of defamation already hinges in important ways on the subject matter of the allegedly libelous speech. See note 168. Moreover, even the inquiry as to whether a plaintiff is a public or private figure may serve as a rough proxy for the more amorphous question whether the case involves speech on matters of public interest. See Elena Kagan, *A Libel Story: Sullivan Then and Now*, 18 L & Soc Inquiry 197, 214-15 & n 39 (1993).

<sup>176</sup> Thomas I. Emerson, *The System of Freedom of Expression* 326 (Random House 1970), quoting *Chaplinsky*, 315 US at 572.

<sup>177</sup> Steven H. Shiffrin, *The First Amendment, Democracy, and Romance* 44 (Harvard

these categories do present a puzzle. Much of the law I have discussed operates to prevent the government from elevating certain forms of speech over others on the basis of their relative value. The government usually cannot say that one kind of speech, as compared with another, is more (or less) worthy, useful, or important. Why, then, may the government say exactly this of fighting words or obscenity? The dearth of public value in such speech may explain the presence of low-value categories, but only by creating a tension within the doctrine of free expression.

A motive-based approach, in contrast, could use a multitiered system in a way consonant with the other central features of First Amendment law. Under this approach, the government would have unusual leeway over so-called low-value speech not because the speech in fact has low value, but because regulation of the speech has a low probability of stemming from illicit motive.<sup>178</sup> If a regulation of a certain kind of speech carries a reduced suspicion of taint, the Court should adopt a standard of review that places a reduced burden of justification on the government; the Court thus should lower the usual strong presumption against regulation or even switch the presumption in the opposite direction. If such a regulation poses no, or almost no, concern of impermissible motive, the Court should allow the legislature to regulate at will, adopting an irrebuttable presumption that the regulation comports with the First Amendment. In this way, so-called low-value categories of speech would appear, but these categories now would fit into, rather than obtrude from, the overall panorama of First Amendment doctrine.

This understanding of a tiered system of expression plausibly explains many, though not all, of the content-based categories of speech that the Court has created. At one end of the spectrum, the regulation of speech about political issues poses the greatest risk of stemming from improper purpose.<sup>179</sup> Political speech often implicates the self-interest of governmental officials; likewise

---

1990).

<sup>178</sup> Cass Sunstein's proposed four-part test of low-value speech includes the inquiry whether "government is unlikely to be acting for constitutionally impermissible reasons . . ." Cass R. Sunstein, *Pornography and the First Amendment*, 1986 Duke L J 589, 604. This matches the approach I am taking; it has nothing to do with the "value" of low-value speech, either to the speaker or to the audience. Sunstein's other three factors do relate to such value. See *id.* at 603-04.

<sup>179</sup> See Thomas I. Emerson, *Toward A General Theory of the First Amendment* 9-10 (Random House 1966); Meiklejohn, *Political Freedom* at 25-27 (cited in note 15); T.M. Scanlon, Jr., *Freedom of Expression and Categories of Expression*, 40 U Pitt L Rev 519, 534-35 (1979).

(and perhaps synonymously), political speech often implicates their most strongly held views and opinions. Because this is so, courts view the regulation of political speech with special disfavor (even if not through a different formal standard), requiring the government to make an extraordinary showing to dissipate the suspicion of improper motive.

At the other end of the spectrum, the regulation of speech falling within low-value categories often raises fewer concerns than usual about improper purpose. The Court's defamation law illustrates the point. The First Amendment provides greatest protection in suits brought by public officials and figures—suits in which the speech most likely addresses matters of public concern and thus most likely implicates the views and interests of decision makers.<sup>180</sup> The First Amendment provides least protection in suits brought by private figures concerning speech of purely private interest—suits in which improper factors are least likely to influence libel judgments.

Similarly, the treatment of commercial speech may respond to the chance that improper motive will taint regulation; in this sphere, the Court may believe, the government less often acts for self-interested or ideological reasons.<sup>181</sup> The government, after all, daily regulates commercial transactions; and in reviewing this regulation, the Court presumes the legitimacy of the government's reasons. A like presumption seems natural when courts review regulation of speech soliciting commercial transactions—itsself a kind of commercial activity. Similar legitimate reasons support such regulation, and the danger of taint appears correspondingly slim. The Court thus lowers the burden placed on the government to demonstrate that regulation of commercial speech has a legitimate basis. Here, officials freely may regulate false and misleading expression—as they also may in the “safe” context of private defamation, but as they may not in the “unsafe” context of political discourse. And as to truthful commercial speech, where a legitimate reason for regulating is not so obvious, officials may act so long as they meet an intermediate stan-

---

<sup>180</sup> See Frederick Schauer, *Public Figures*, 25 *Wm & Mary L. Rev.* 905, 921-29 (1984). Schauer argues that fear of governmental overreaching explains the treatment of libel suits brought by public officials, but not by public figures. But if the inquiry into whether the plaintiff is a public figure represents a more easily administrable version of the inquiry into whether the speech concerns a matter of public interest, see note 175, then the fear that impermissible factors will influence judges or jurors also explains the cases involving public figures.

<sup>181</sup> See Scanlon, 40 *U Pitt L. Rev.* at 541 (cited in note 179).

dard. The rigor of review again follows the danger that illicit motive has tainted a law.

The analysis proceeds in a slightly different way for categories of speech entitled to no protection. In creating these no-value categories, the Court may be stating an irrebuttable presumption that certain prohibitions arise from legitimate motives. Here, the Court in effect predecides that a ban on all speech within the category will result not from hostility toward the ideas restricted, nor from self-interest, but instead from a neutral decision that the harm the speech causes justifies prohibition. This predecision amounts to much the same thing as a post hoc decision in the usual case that a regulation meets the requisite standard of review. Both are conclusions about the absence of illicit motive; the difference is only that in the former case the constitutional review and the conclusion as to motive are submerged in the creation of an "unprotected" category. What the Court is saying is that any law framed in the way the Court has defined the category closely enough fits a sufficiently significant interest as to negate any worry of improper purpose.

This explanation fits the Court's creation of a fighting words category. Though hazy in its boundaries, this category seems to embrace direct face-to-face insults that would cause the average addressee to respond with violence. In holding that a legislature may prohibit fighting words, the Court is doing no more than approving a governmental response to an immediate danger of violence. The premise of the category, no less than of a clear-and-present-danger test, is that the government would respond to such a danger no matter what its views of the ideas affected.

A motive-based account, however, fits poorly if at all the Court's treatment of obscene speech. As an initial matter, the formal test for determining obscenity suggests that motive is not the key. That test mandates an inquiry into the value of the materials, which suggests that the concerns of the audience-based model here predominate. The test also demands a finding of community offense, which constitutes, rather than disproves, improper motive. And even if the formal test did not include these attributes, the probability of taint infecting an obscenity law seems severe. Such a law might stem from a neutral evaluation of the harm these materials cause, such as sexual violence. But hostility toward certain ideas about sexual mores—otherwise stated, the desire to maintain status quo ideas about sexuality free from challenge—are likely to color this evaluation or trump it entirely. (Consider how officials often respond to speech show-

ing or approving homosexuality.) In this area, a motive-based model thus fails to explain the doctrine.<sup>182</sup>

But placing sexually explicit speech to one side, the term "low-value" category may be a misnomer. Perhaps what sets these categories apart is not that the speech within them is low value, but that regulation of the speech within them is low risk. No matter that a regulation of these categories is content based, even viewpoint based; the government need not satisfy the usual standard because the courts do not suspect, to the usual extent, that the government's asserted legitimate interest is a pretext. The categories—call them now "low-risk categories"—thus function as yet another part of the broad evidentiary mechanism designed to flush out actions based on impermissible purpose.

This explanation makes the holding of *R.A.V.*—invalidating a partial prohibition of fighting words—consonant with the theory that underlies the fighting words category.<sup>183</sup> This category rests on a presumption that, given the state's great and legitimate interest in maintaining order, a ban on fighting words stems from proper motive. The Court has no reason to apply this presumption when the government regulates only part of the category, on the basis of content (or viewpoint) extraneous to the category's boundaries.<sup>184</sup> In that case, there is an underinclusion problem—a lack of fit between the regulation and

---

<sup>182</sup> The motive-based model also may not explain the category of child pornography, though here the analysis is more complicated. The formal test for child pornography does not include any assessment of the material's offensiveness or value. See *Ferber*, 458 US at 754-65. In addition, the harms caused by this material to children would seem amply to justify governmental action. But here too, motives are complex, and it would be surprising if disapproval of certain ideas about sexuality and children did not enter into the government's calculation.

<sup>183</sup> See *Kagan*, 60 U Chi L Rev at 899-900 (cited in note 61).

<sup>184</sup> The situation would be different, as Justice Scalia noted in *R.A.V.*, if the government regulated only part of the category, but on the basis of the characteristics (and corresponding state interests) that define the category. 505 US at 387-88. Justice Scalia used the example of a law that prohibited only the most prurient of all obscenity. *Id.* at 388. In his words:

When the basis for the content discrimination consists entirely of the very reason the entire class of speech at issue is proscribable, no significant danger of idea . . . discrimination exists. Such a reason, having been adjudged neutral enough to support exclusion of the entire class of speech from First Amendment protection, is also neutral enough to form the basis of distinction within the class.

Justice Scalia's conclusion that the St. Paul law did not fit within the contours of this principle is open to question. See *id.* at 403-04 (White concurring); *id.* at 422-24 (Stevens concurring). But his broader point is on the mark: it understands the critical issue, in evaluating regulation, as relating to the government's motive.

the reasons supporting the fighting words category. This lack of fit provides grounds to suspect that the government is acting for other reasons, which the Court has yet to identify, let alone find legitimate. So whereas a full-scale ban on fighting words is "low risk," a partial, content-based ban on fighting words is not. The Court's treatment of each relates to the probability of taint, but that single gauge points in opposite directions.

A second recent case—*City of Cincinnati v Discovery Network, Inc.*<sup>185</sup>—also highlights the connection between governmental motive and so-called low-value categories. The case began when city officials banned newsracks disseminating commercial handbills—62 of the approximately 1,500 newsracks cluttering Cincinnati's sidewalks. The city asserted, as its basis for action, interests in safety and aesthetics. Of course, commercial and noncommercial newsracks look the same and thus implicate these interests in the same way. The city noted, however, that commercial speech had "low value."<sup>186</sup> If this were true, the city claimed, then its interests could justify a ban on commercial newsracks, but not any others. This claim, accepting the premise that commercial speech has low value, is unexceptionable; and the Court accepts exactly this premise in commercial speech cases. Yet the Court rejected the city's argument. Why?

Said the Court: Because the distinction between commercial and noncommercial speech "bears no relationship *whatsoever* to the particular interests [in safety and aesthetics] that the city has asserted."<sup>187</sup> Well, yes—the city had admitted as much. But why should this matter if commercial and noncommercial speech have different values, such that a single interest can outweigh the one kind of speech but not the other? Said the Court: "In our view, the city's argument attached more importance to the distinction between commercial and noncommercial speech than our cases warrant and seriously underestimates the value of commercial speech."<sup>188</sup> Did this mean commercial speech henceforth would receive full constitutional protection? In cases after *Discovery Network*, the Court has continued to treat such speech as falling within a low-value category.<sup>189</sup> And in *Discovery Network*

<sup>185</sup> 507 US 410 (1993).

<sup>186</sup> *Id.* at 418-19.

<sup>187</sup> *Id.* at 424. See also *id.* at 426-31.

<sup>188</sup> *Id.* at 419. See also *id.* at 428-31.

<sup>189</sup> See *Rubin v Coors Brewing Co.*, 115 S Ct 1585, 1589 (1995).

itself, the Court applied not strict scrutiny, but the intermediate standard of review devised for commercial speech cases.

The only way to reconcile *Discovery Network* with the Court's commercial speech doctrine is to see that doctrine as emerging from a judgment of risk, rather than a judgment of value. Once the doctrine is viewed in this way, the Court can apply relaxed scrutiny to regulations of commercial speech at the same time as it instructs governmental officials to treat this kind of speech as having the worth of any other. The problem in *Discovery Network* was that Cincinnati had mistaken the basis of the commercial speech category. Cincinnati had treated the creation of this category as a kind of substantive statement, addressed to courts and legislatures alike. The city had argued: Commercial speech has low value and therefore may be regulated on a reduced finding of harm. The Court in *Discovery Network* responded that the creation of the category was instead an evidentiary statement, addressed only to courts. The Court, that is, had said: Regulation of commercial speech has low risk and therefore should be held to a reduced standard of review. *Discovery Network* thus well illustrates the true nature—contrary to the appellation—of low-value categories. On this understanding, low-value categories fall into line with the rest of First Amendment law; they become another way of focusing and refining the search for motive.

## 2. Secondary effects doctrine.

The Court in recent years has toyed with an approach, known as the "secondary effects" doctrine, that would treat another swath of content-based regulations as if content neutral. This doctrine, in focusing on the nature of governmental justification, appears at first glance to prove my theory. The reality is more complex. The motive-based model and the secondary effects doctrine operate in considerable tension with each other. Even so, this model comes closer than its competitors to explaining the doctrine of secondary effects. This explanation turns, once again, on the relative "safety," with respect to motive, of certain kinds of content-based actions.

The essence of the secondary effects doctrine runs as follows: facially content-based regulations of speech that "are *justified* without reference to the content of the regulated speech" should be treated as if they made no facial distinctions on the basis of content.<sup>190</sup> In the seminal case, *City of Renton v Playtime The-*

---

<sup>190</sup> *City of Renton v Playtime Theatres, Inc.*, 475 US 41, 48 (1986), quoting *Virginia*

aters, the Court approved, under the standard of review usually used to test facially content-neutral action, a zoning ordinance applying to (and only to) theaters that showed sexually explicit movies.<sup>191</sup> The Court explained its choice of standard by reference to the aim of the zoning ordinance. The ordinance, the Court emphasized, was not intended to "suppress the expression of unpopular views"<sup>192</sup> or to "restrict[ ] the message purveyed by adult theaters."<sup>193</sup> Rather, the law was "designed" to achieve certain so-called secondary effects—specifically, "to prevent crime, protect the city's retail trade, maintain property values, and generally 'protect and preserv[e] the quality of [the city's] neighborhoods.'"<sup>194</sup> Because the ordinance aimed at these objects, it did not demand heightened scrutiny.

This doctrine contravenes the strictures of both the speaker-based and the audience-based models of the First Amendment. Consider two laws, one restricting sexually explicit speech to preserve neighborhood character, the other restricting the identical speech to suppress the message of sexual libertinism. The two laws curtail to the same extent both a speaker's opportunity to express a message and a listener's opportunity to consider its merits. Under each model, then, the difference between these two laws should have no constitutional significance. The secondary effects doctrine makes that difference relevant because, unlike these effects-based models and despite its name, it concerns itself with a regulation's reasons.

Stated in this way, the secondary effects doctrine seems to offer a particularly fine example of the way in which the search for motive structures the law of the First Amendment. Indeed, an understanding of this kind might explain the Court's creation of the doctrine. To the Court, the doctrine may have seemed but a natural—even an essential—aspect of First Amendment law's focus on reasons.

In fact, however, the secondary effects doctrine fits uneasily with the rest of First Amendment jurisprudence. Although the

---

*Pharmacy Bd v Virginia Citizens Consumer Council, Inc.*, 425 US 748, 771 (1976) (emphasis in *Renton*). The Court has not yet determined whether the secondary effects doctrine applies in the realm of fully protected speech. The Court has used the doctrine only when dealing with sexually explicit expression. The Court, however, has not made the presence of arguably low-value speech a definite condition of the doctrine's application and once has suggested to the contrary. See *Boos v Barry*, 475 US 312, 320-21 (1988).

<sup>191</sup> 475 US 41, 54-55 (1986).

<sup>192</sup> *Id.* at 48.

<sup>193</sup> *Id.*, quoting *Young v American Mini Theatres, Inc.*, 427 US 50, 82 n 4 (1976).

<sup>194</sup> *Id.*, quoting the appendix to the Jurisdictional Statement.

doctrine reflects the First Amendment's concern with illicit motive, it disdains the usual method by which that concern is effected. Rather than using objective tests as proxies for examining motive, the doctrine of secondary effects insists on directly addressing this question. Construed in one way, the doctrine requires courts to make independent determinations of the government's reasons, notwithstanding the difficulties of proof entailed in this effort. Construed in another, more practical way, the doctrine counsels courts to accept without investigation plausible assertions as to governmental motive. Under either approach, courts often will countenance what they should not, mistaking the pretextual reason for the real and the tainted motive for the pure. The secondary effects doctrine thus seems to run counter to—and thereby negate the effect of—all the indirect techniques for flushing out illicit purpose that the Court has developed.

Is, then, the secondary effects doctrine a simple mistake—a doctrine meant to address the danger of improper purpose, but designed so as to undermine its own objective? Before reaching this conclusion, a closer examination of the doctrine is in order. Such an inspection may show the secondary effects doctrine to be a more sensible response to the fear of improper purpose than the I have just suggested. Or, less optimistically, this investigation at least may suggest ways of interpreting and shaping the doctrine so as to make it an aid, rather than a hindrance, in the effort to detect improper motive.

The distinction that drives the secondary effects doctrine, at least in its most sensible incarnation, is a distinction between the communicative (primary) impact and the noncommunicative (secondary) impact of expression. An example focusing on the regulation of child pornography will illustrate the point. Assume that two municipalities enact an identical child pornography statute. City A does so on the ground that viewing child pornography increases the probability that an individual will commit an act of child molestation. City B does so on the ground that manufacturing child pornography itself involves acts of child molestation. Which statute (both, either, none) should the Court treat as content neutral because based on secondary effects?

On one conceivable view of the secondary effects doctrine, both statutes would fall within the doctrine's scope. Neither statute, after all, arises from a censorial purpose—from the simple desire, in the words of *Renton*, to “suppress expression of unpopu-

lar views."<sup>195</sup> Both statutes instead arise from the desire to prevent real-world harms of a serious nature; hence, both provoke application of the doctrine of secondary effects. But this understanding of the secondary effects doctrine would revolutionize free speech law. It would turn every First Amendment case into a secondary effects case, given that the government almost always can proffer a justification based on harm.

A second view of the secondary effects doctrine avoids this outcome by narrowing the doctrine's scope. Under this view, City B's justification for regulating child pornography is based on a secondary effect of the expression, but City A's is not. The difference lies in whether the harm the government is seeking to prevent arises from the expressive aspect of the communication—or, stated in another way, whether the harm results from a listener's hearing the content of speech and reacting to it. The harm justifying City A's ordinance arises from such a communicative effect: the content of the speech moves a listener to engage in hurtful behavior. The harm justifying City B's ordinance does not arise from the content of communication in this way. Although the harm relates to expression—more, to expression of a certain content—the harm does not grow out of a listener's response to a message. In the language of three Justices who appear now to represent the view of the Court on this issue, the government justifies its regulation of content on the ground that "the regulatory targets happen to be associated with" this content, rather than on the ground that the comixture between the content and the listener itself brings the regulatory target into being.<sup>196</sup>

---

<sup>195</sup> *Renton*, 475 US at 48. Chief Justice Rehnquist, the author of *Renton*, may well hold this broad view of the secondary effects doctrine. All of the language in *Renton* distinguishes between a desire to suppress speech because of the harm it causes and a desire to suppress speech because of the ideas it espouses. By contrast, none of the language distinguishes between harms arising from the communicative aspects of speech and harms arising from the noncommunicative aspects. See *id.* at 47-49. Indeed, the Chief Justice, along with several other Justices, refused to join a portion of a later opinion attempting to narrow application of the secondary effects doctrine to regulations of speech based on noncommunicative impact. See *Boos*, 485 US at 338 (Rehnquist concurring).

<sup>196</sup> *Boos*, 485 US at 320 (O'Connor opinion). Secondary and primary effects, defined in this way, are sometimes difficult to tell apart. *Renton* provides one example: is increased crime merely "associated with" sexually explicit speech, or does it result from the effects of such speech on an audience? Or consider a statute drawing a content-based distinction on the ground that speech of a certain content draws large crowds and thus causes congestion. Does the congestion arise from the communicative effect of the speech or is the congestion merely associated with the expression? Many similar quandaries could be devised. But in this context, as in many others, the ability to concoct cases that might fit into one or another category does not gainsay the usefulness of the categories generally. In most instances, the distinction between primary and secondary effects, drawn in the

Why should this distinction between communicative and noncommunicative impact count as crucial in delimiting the doctrine of secondary effects? On one explanation, the distinction matters because it marks the precise divide between permissible and impermissible motive in the regulation of expression. The government, that is, may not restrict expression for any reason relating to communicative impact; the government may limit speech only for reasons independent of the response of listeners to a message.<sup>197</sup> Thus follows naturally the role of communicative impact in shaping the doctrine of secondary effects: the Court, in asking whether a restriction is justified in terms of communicative impact, is doing no more and no less than asking whether a restriction is tainted with improper motive.

But this explanation is unsatisfying. I have argued before that the concept of impermissible purpose posited in this account conflates dissimilar governmental motives and conflicts with the Court's First Amendment decisions.<sup>198</sup> More important for current purposes, the explanation just offered fails to address, much less to dissolve, the criticism that the doctrine of secondary effects hinders, rather than aids, the effort to uncover improper purpose. If all reasons relating to the communicative impact of speech are improper, then the current doctrine of secondary effects indeed turns on the concept of illicit motive. But in that case, the doctrine defeats its own mission by discarding the (necessary) indirect means of discovering motive and substituting a (hopeless) direct inquiry.

A second and better explanation for making communicative impact the key to the doctrine of secondary effects is that the resulting doctrine refines courts' efforts to flush out the narrow class of motives I have labeled improper (not all reasons relating to communicative impact, but only reasons of ideology or self-interest). On this view, the concept of communicative impact, instead of defining what counts as legitimate purpose, plays a quasi-evidentiary role, signaling a change in the standard of review a court needs to uncover the presence of improper motive. Assume here that improper motive, as I have defined it, is easier to detect—and less likely, in the first instance, to exist—when the justification for a statute relates to noncommunicative, rather

---

way described in the text, can be applied in a sensible manner.

<sup>197</sup> I have discussed this view of impermissible purpose, suggested in the writings of John Hart Ely and Laurence Tribe, at text accompanying notes 67-71.

<sup>198</sup> See text accompanying notes 60-66, 69-71.

than communicative, impact. In this event, a relaxed standard of review would suffice in secondary effects cases to separate proper and improper motives, and a court should discard its usual "sledgehammer" standard for a daintier constitutional instrument. The secondary effects doctrine, it might be argued, accomplishes just this result, thus rendering still more precise the complex mechanism established by First Amendment law to uncover improper motive.

This theory, of course, rests on a nonobvious premise: that courts can detect improper purpose more easily, and legislators will resort to it less often, when the justification for a restriction refers to secondary, rather than primary, effects. Why would this be? The key point is that because the harm in secondary effects cases derives from a thing only contingently related to expression, courts and legislators in these cases possess, to a greater degree than is usual, two testing devices for stripping away pretexts and revealing motives.

First, a court usually can check for improper motive in a secondary effects case by asking whether the government has tried to regulate the affected speech in the absence of the harm asserted. Suppose, as in *Discovery Network*, that the government justifies a restriction on commercial publications by pointing to the aesthetic deficiencies of newsracks; then the restriction can go no further than to limit this one means of distribution.<sup>199</sup> Or suppose, as in *Simon & Schuster, Inc. v Members of the New York Crime Victims Bd*, that the government justifies a law impounding profits from published descriptions of crime as a way to compensate the crime's victims; then the provision can apply only when "a victim remains uncompensated."<sup>200</sup> These examples suggest that secondary effects cases often involve narrower restrictions on speech than do primary effects cases.<sup>201</sup> But the critical point concerns not the breadth of the restriction, but the ease of deciding whether the justification offered for it is real or pretextual. In almost any secondary effects case, a control group

---

<sup>199</sup> 507 US at 425-26. The restriction at issue was so limited.

<sup>200</sup> 502 US 105, 122 (1991). The Court found the statute overinclusive in part because it applied to works that did not implicate the interest in compensation. *Id.*

<sup>201</sup> See Scanlon, 40 U Pitt L Rev at 528 (cited in note 179). This is usually, but not necessarily, true. A flat prohibition on speech may derive from a secondary justification. See, for example, *New York v Ferber*, 458 US 747 (1982) (ban on child pornography based on harm to the child photographed). Conversely, a limited regulation may derive from a primary justification. See, for example, *Nebraska Press Association v Stuart*, 427 US 539 (1976) (restrictions on press coverage of trial based on potential response of public).

will exist consisting of the identical speech unaccompanied by the (nonexpressive) agent causing the asserted harm. A check of this kind will exist more rarely in primary effects cases because there the harm derives from the substance of the speech, rather than from features severable from it.<sup>202</sup> That difference may make the difference in the effort to flush out improper motive.

Second, a court usually can check for improper motive in a secondary effects case by asking whether the government has regulated conduct, as well as other speech, that causes the same harm as the affected expression. The Court, for example, in *Discovery Network* asked whether the government had prohibited newsracks not only for commercial, but also for noncommercial publications, given that the two posed identical aesthetic issues.<sup>203</sup> Similarly, in *Simon & Schuster*, the Court asked why the statute covered only the profits criminals gained from expressive activity, when the profits gained from nonexpressive activity also could have compensated victims of crime.<sup>204</sup> Again, the ability to check the government's justification in this way depends in part on whether the case involves primary or secondary effects. If the government invokes primary (communicative) effects, then a court often (though not always) will have no nonspeech analogues by which to test the restriction; nonexpressive activity, or even expressive activity of a different content, will rarely implicate in the same way the same governmental interest. But if the government invokes secondary (non-communicative) effects, then a court usually (though again not always) will have a reference point in the government's treatment of nonexpressive or other expressive activity; this is so precisely because the asserted harm in the case comes from a non-communicative, hence severable, aspect of the expression.

Because of the common availability of these testing mechanisms in secondary effects cases, courts arguably do not need to apply the most stringent standard of review in order to flush out improper motive. Even the standard of review applicable to con-

---

<sup>202</sup> There are exceptions to each side of this comparison. First, if a noncommunicative aspect of speech attaches to all speech of a certain content, then the control group I am positing in secondary effects cases will not exist. I doubt, however, that there are many cases of this kind. Second, if a communicative aspect of speech causes harm only in particular circumstances, then a control group might exist in primary effects cases consisting of the same speech occurring in all other circumstances. I doubt, however, that such a verification device will arise as often or prove as easy to apply as the checking device available in secondary effects cases.

<sup>203</sup> See 507 US at 425-26.

<sup>204</sup> See 502 US at 119-20.

tent-neutral action bars any law that (1) restricts speech not implicating the governmental interests asserted, or (2) discriminates against speech relative to nonexpressive activities implicating the same interests.<sup>205</sup> If the testing mechanisms I have discussed make this standard an effective filter in secondary effects cases, then use of a stricter standard would be overkill. When the nature of the asserted justification makes it easy to impeach, then courts do not need the powerful weapons designed to achieve this object in other circumstances.

I do not wish to stake very much on the strength of this motive-based explanation of the doctrine of secondary effects. Perhaps I have overstated the difference between primary and secondary effects cases with respect to the difficulty of discrediting pretexts and finding real reasons. If so, the doctrine of secondary effects constitutes at once a paradox and an error. The paradox is that the doctrine of First Amendment law most concerned with evaluating reasons is the doctrine least reconcilable with the motive-based model. The error, which creates the paradox just mentioned, lies in the decision to evaluate reasons by asking questions about them. In fact, the proper way to resolve the issue of motive is to pose other questions. If the secondary effects doctrine indeed is incompatible with the motive-based approach, it is because the former brings to the surface what the latter knows should reside beneath it.

But if the doctrine of secondary effects has any sound foundation, it relates to refining the search for improperly motivated governmental actions. More specifically, the doctrine emerges from the view that it is relatively easy in cases involving secondary effects to isolate the role played by hostility, sympathy, or self-interest. No other account of the doctrine of secondary effects makes better (or indeed, any) sense. And this account provides a guide for judges applying the doctrine. In highlighting the way in which a secondary justification affects the search for improper

---

<sup>205</sup> The standard applicable to content-neutral regulations, although variously articulated, requires that a law extend only to speech implicating the asserted interest. See *Ward v Rock Against Racism*, 491 US 781, 798 (1989) (requiring that a regulation "be narrowly tailored to serve the government's legitimate content-neutral interests"); *O'Brien*, 391 US at 377 (demanding that a restriction be "no greater than is essential to the furtherance of [the asserted] interest"). Although not included in statements of the applicable standard of review, the principle that the government may not enact content-neutral laws discriminating against speech, as compared to nonexpressive activity, appears in (and alone explains) many cases. See *Martin v City of Struthers*, 319 US 141, 142-44 (1943); *Schad v Borough of Mount Ephraim*, 452 US 61, 73 (1981); *Minneapolis Star & Tribune Co. v Minnesota Comm'r of Revenue*, 460 US 575, 581-85 (1983).

motive, the account focuses judicial attention on that search, while revealing the checking devices available to conduct it. The doctrine of secondary effects thus adds further intricacy to the essential mechanism of First Amendment law: a mechanism devised to flush out improper purpose on the part of the government in the face of serious, but variable, problems of proof.

#### D. The Distinction Between Direct and Incidental Restrictions

Consider again the facts of *R.A.V.*, but this time with a twist, simple in form yet profound in consequence. Recall that in the actual case, St. Paul charged a juvenile who allegedly had burned a cross on the property of an African-American family with violating the city's hate-speech ordinance, a law interpreted to punish, and to punish only, certain kinds of expressive activity. Now suppose the conduct of the juvenile remains the same, but the prosecutor makes a different charging decision. Rather than resort to the city's hate-speech ordinance, the prosecutor relies on a generally applicable law—a law not specifically directed toward speech or other expressive activity. Any of a number of laws may come to mind—a statute prohibiting trespass or arson or the infliction of damage on another person's property.<sup>206</sup> If the prosecution had proceeded under one of these laws—say, the trespass ordinance—the trial court would have dismissed a defense based on the First Amendment. Even assuming the burning of a cross qualifies as fully protected expressive activity, the case would have started and ended as a trespass prosecution.

This result arises from a distinction as important as any in First Amendment law: the distinction between direct and incidental restrictions on speech or, otherwise phrased, the distinction between actions targeting expression alone and actions applying generally, to both nonexpressive and expressive activity.<sup>207</sup> The distinction received its most famous articulation in

---

<sup>206</sup> The Court in *R.A.V.* noted the possibility that the conduct at issue violated statutes of this kind, thus implying that a prosecution brought under such a statute would have raised no First Amendment problems. 505 US at 380 n 1.

<sup>207</sup> See generally Frederick Schauer, *Cuban Cigars, Cuban Books, and the Problem of Incidental Restrictions on Communications*, 26 Wm & Mary L Rev 779 (1985); Stone, 54 U Chi L Rev at 105-14 (cited in note 110). This distinction, of course, assumes another: between expressive and nonexpressive activity, sometimes loosely termed speech and conduct. I have noted elsewhere the need for First Amendment law to view expressive and nonexpressive activity as meaningfully different, even though drawing a line between the two raises hard questions. See Kagan, 60 U Chi L Rev at 883-84 (cited in note 61). For purposes of this Article, the more important point is that the Court always has distinguished between the two, usually by asking whether the activity in question is, in purpose

*O'Brien*, when the Court upheld the conviction of an antiwar protester who had burned his draft card for violating a federal statute prohibiting the intentional destruction of draft registration certificates. The Court explained that the prosecution comported with the First Amendment because the statute (and the governmental interest supporting it) applied regardless whether the conduct enjoined was expressive; hence the restriction on speech was incidental.<sup>208</sup> As the Court more recently stated the proposition in *R.A.V.*, "[w]here the government does not target conduct on the basis of its expressive content, acts are not shielded from regulation merely because they express a[n] [ ] idea or philosophy."<sup>209</sup> So long as the law applies generally, it can apply to expression. Thus the distinction between prosecuting a cross burner under a trespass law and prosecuting him under a hate-speech ordinance.

Courts usually treat the application of a general law, even to activity concededly expressive, as raising no First Amendment issue whatsoever. So, for example, courts will not see a constitutional question if the government convicts for vandalism a person who draws swastikas on a synagogue wall; or applies taxation, labor, or antitrust laws to the publisher of a newspaper; or uses a residential zoning law to prevent the opening of a bookstore. On occasion, courts will apply a form of intermediate scrutiny to an incidental restraint, as the Court did in *O'Brien*,<sup>210</sup> still more rarely, a court will subject to strict scrutiny the use of a generally applicable statute.<sup>211</sup> I will have more to say later about these more or less exceptional cases. But for now, the critical point is that incidental restraints on expression usually receive more deferential treatment than direct restraints on the same expression.<sup>212</sup> The question is why this is so.

---

and function, primarily expressive. See *Spence v Washington*, 418 US 405, 410-11 (1974) (asking about the "intent to convey a particularized message" and the likelihood that "the message would be understood by those who viewed it"). I use this view of the difference between expressive and nonexpressive activity in this discussion.

<sup>208</sup> *O'Brien*, 391 US at 376-82. Earlier statements of this principle appear in *Associated Press v NLRB*, 301 US 103, 132 (1937) (applying labor laws to press), and *Associated Press v United States*, 326 US 1, 7 (1945) (applying Sherman Act to press).

<sup>209</sup> 505 US at 390.

<sup>210</sup> 391 US at 376-77. See text accompanying notes 225-34.

<sup>211</sup> See *United States v Eichman*, 496 US 310, 318 (1990). See also text accompanying notes 235-43.

<sup>212</sup> It is on account of this principle—and this principle alone—that the Court is right to say that "[t]he Government generally has a freer hand in restricting expressive conduct than it has in restricting the written or spoken word." *Texas v Johnson*, 491 US 397, 406 (1989). The "freer hand" is a function not of any difference between verbal and nonverbal

If the law concerning direct and incidental restraints seems too obvious to merit this discussion, consider as a preliminary matter that the distinction first arose, and then operated for years, in a form converse to that of modern doctrine. In *Gitlow v New York*, the Court contrasted a statute directly targeting certain forms of advocacy to a statute "prohibit[ing] certain acts involving the danger of substantive evil, without any reference to language itself."<sup>213</sup> The Court noted that it had greater call to review an application of the general statute to speech than to review an application of the targeted statute. In the latter case, the Court seemed to reason, the legislature already had made a considered judgment that the speech at issue posed the requisite danger; in the former case, the legislature had made no such judgment, and might have concluded to the contrary, had it ever considered the matter.<sup>214</sup> The unintentionality of the incidental restraint worked against it; the purposefulness of the direct restraint worked in its favor. The end result was a doctrine that treated the restriction of speech through a generally applicable law as more, rather than less, problematic than the restriction of speech through direct legislation.

What accounted for the Court's eventual shift in understanding? Again, I argue that modern doctrine acquired its structure as an attempt to discover actions tainted with ideological motive. What changed between the old case law and the new was the Court's understanding of its role in policing the inputs of governmental action affecting expression. Before I press this point, however, I again consider alternative arguments.

The inadequacy of a speaker-based model in explaining current doctrine emerges from a simple form of hypothetical, contrasting a direct restriction on speech to a generally applicable regulation that covers everything affected by the direct restriction, but more in addition. Suppose, for example, that one city makes it illegal to deface synagogues with swastikas, while another enjoins all acts of vandalism. Or that one city imposes a tax on corporations that publish newspapers, while another imposes the tax on corporations generally. Or, finally, that one city prohibits bookstores in an area, while another zones the area for

---

expression, but of the difference between laws targeting expression and laws applying more broadly; the government may restrict expressive conduct more easily than the written or spoken word because expressive conduct more often falls within the terms of generally applicable regulations.

<sup>213</sup> 268 US 652, 670 (1925).

<sup>214</sup> *Id.* at 670-71.

residential use only.<sup>215</sup> If what mattered were the effect of a regulation on a speaker's expressive opportunities, then courts would review these municipal acts in identical fashion; after all, the one city impairs speech no more than the other.<sup>216</sup> Of course, the generally applicable law may have a stronger state interest supporting it than does the targeted restriction. But this difference, even when it exists, would not justify separate standards of review; at most, it would explain how application of the same standard to these laws could produce different outcomes. The current doctrinal structure thus cannot result from an inquiry into the effects of a law on expressive opportunities.

Neither can the distinction between direct and incidental restrictions derive from a concern with the way in which governmental actions distort public discourse. Even supposing that incidental restrictions were subject to the same level of review as other content-neutral laws, this account of the doctrine would fail to persuade. It then would suffer from all the same difficulties as plague the attempt to explain, on grounds of distortion, the distinction between content-based and content-neutral laws generally.<sup>217</sup> Most notably, the account would ignore that any incidental restraint works against a backdrop of other laws—themselves functioning as incidental restraints—that may render the restraint a means of perpetuating not balance in discourse, but distortion. That laws affecting speech “only” incidentally may play a large role in shaping the speech market, either for good or for ill, should not by now be in question.

But using an audience-based model here poses further problems, for it cannot explain why current doctrine in fact treats incidental restrictions more deferentially than other (speech-specific) content-neutral restrictions. Even assuming that a law distorts public discourse to the extent it alters the *ex ante* distribution of opinion,<sup>218</sup> there is no reason to think that content-neutral direct restraints cause greater distortion than incidental

---

<sup>215</sup> The last example comes from Stone, 54 U Chi L Rev at 105 (cited in note 110). See also Schauer, 26 Wm & Mary L Rev at 779 (cited in note 207) (listing similar examples).

<sup>216</sup> Indeed, the city using general restrictions may curtail more speech than the city using direct restrictions. The general tax, for example, covers broadcast stations as well as newspapers, and the general zoning law covers not only bookstores, but video outlets.

<sup>217</sup> See text accompanying notes 91-100.

<sup>218</sup> I am indulging this assumption, which is contrary to all prior argument, to cast in the best possible light the audience-based explanation of the doctrine of incidental restraints. If the assumption is dropped, the explanation will fail regardless whether generally applicable laws or targeted content-neutral laws more greatly alter the existing distribution of opinion.

restrictions. Indeed, if either were to skew debate more than the other, incidental restrictions would count as the culprit. Whereas the defining characteristic of content-neutral direct restrictions is what I will call "horizontal sweep" (application to different kinds of speech), the defining characteristic of incidental restraints is "vertical sweep" (application to speech and conduct), which may or may not have a horizontal dimension. Consider, for example, the law in *O'Brien*, which in prohibiting the destruction of draft cards interfered with one viewpoint only. Given such cases, an approach focusing on distortion might turn current doctrine inside-out and require courts to review incidental restraints more closely than direct restrictions.<sup>219</sup>

Perhaps the explanation of current doctrine lies solely in a set of practical constraints.<sup>220</sup> If all laws incidentally restricting speech were subject to First Amendment review, then (almost) all laws would be subject to First Amendment review. This is to say no more than what I have suggested before: that many laws not specifically directed at speech have effects on expressive activity. The need to address the constitutionality of all such laws would impose significant costs. If, as seems likely, most of the laws would pass constitutional muster, incurring these costs does not seem worthwhile.<sup>221</sup> Better to assume from the outset that these laws raise no serious problem.

But an alternative or additional rationale is available: that the law of incidental restraints arises from a focus on governmental motive.<sup>222</sup> Consider that a generally applicable law by definition targets not a particular idea, nor even ideas broadly speaking, but an object that need not, and usually does not, have any association with ideas whatsoever. Recall the examples used

---

<sup>219</sup> I discuss below why a motive-based account would not do so, even given that differential effects are one indicator of improper motive. See text accompanying note 223.

<sup>220</sup> See Schauer, 26 *Wm & Mary L Rev* at 784 (cited in note 207); Stone, 54 *U Chi L Rev* at 107 (cited in note 110).

<sup>221</sup> If any significant portion of the laws would fail constitutional scrutiny, a different—but perhaps no less trenchant—objection to judicial review would arise. The core issue then would concern the propriety of interpreting the First Amendment to impose substantial limits on the government's power to enact or apply regulations not specifically addressed to expression. Resolution of this issue would depend on the selection of a theory of the First Amendment; the theory posited in this Article, for reasons that follow, holds that the First Amendment should not be understood to impose these limits.

<sup>222</sup> See Schauer, 26 *Wm & Mary L Rev* at 783 (cited in note 207); Schauer, *Free speech* at 100-01 (cited in note 39). The doctrine, viewed in this way, has an analogy in equal protection law, where the Court has held that facially race-neutral laws pose a relatively minor risk of stemming from impermissible motive and hence should receive relaxed scrutiny. See *Washington v Davis*, 426 US 229, 247-48 (1976).

above: laws prohibiting vandalism, imposing a tax, or setting a zoning restriction. The breadth of these laws makes them poor vehicles for censorial designs; they are instruments too blunt for effecting, or even reflecting, ideological disapproval. Thus, incidental restrictions receive minimal scrutiny because of the likelihood that they also will be accidental restrictions in the relevant sense—that they will result from a process in which officials' hostility (or partiality) toward ideas played no role.

This analysis explains why the Court treats incidental restraints more deferentially even than direct restraints of a content-neutral nature. It is true that the breadth of content-neutral direct restraints also lessens the likelihood of illicit motive. Indeed, to the extent that effects constitute evidence of intent, the argument I made earlier about horizontal and vertical sweep might support not the laxer, but the harsher treatment of incidental restraints than of content-neutral direct restrictions. But the difference in the facial terms of the two kinds of regulations—that, by definition, one goes to speech, the other to conduct—cuts sharply in the direction of existing doctrine.<sup>223</sup> When a proposed law is addressed to expression, a legislator cannot help but consider, consciously or not, whether and how the law will affect particular messages; this is to say little more than that when a law is about speech, the legislator will consider its impact on speech—a proposition neither deep nor shocking. But when a proposed law, by its terms, focuses on nonexpressive conduct, restricting speech only as an incidental and thus a covert matter, the probability increases that a legislator will consider the regulation divorced from hostility or sympathy toward particular messages.

A law of general application, of course, may have such dramatic—and apparent—effects on expressive activities that it might as well target those activities in express terms. I soon note that it is in just these circumstances that courts treat generally applicable laws as if they were direct restrictions on expression. But even were this escape hatch unavailable, a decision to treat incidental restraints with a degree of deference not given to content-neutral direct restrictions would have an adequate justification. If it is not always true, it is true often enough that bias toward ideas will infect the former less than the latter.

---

<sup>223</sup> I similarly argued in Section III.A that the facial terms of a regulation—there, whether the regulation was content based or content neutral—provides a better indication of governmental motive than does any measurement of content-differential effects.

The Court's abandonment of *Gitlow* thus results from its adoption of the foundational principle of modern doctrine. The thing that *Gitlow* counted in a law's favor—the purposeful and considered judgment of the desirability of repressing expression—now appears a grave danger, because it so easily may become tainted with ideological factors. The thing that *Gitlow* counted against a law—the nonobviousness of its relation to, or effects on, expression—now appears a great boon, because when legislators do not consider the question of restricting speech, *a fortiori* they do not consider it in an improper manner. The key point concerns the relation of general applicability to neutrality and of neutrality to motive.<sup>224</sup> Generality implies neutrality among ideas as such; neutrality of this kind signifies and, indeed, defines the absence of impermissible motive.

If, in this way, a concern with motive accounts for the doctrine of incidental restraints, then the same concern ought to aid in explaining exceptions to this doctrine: the cases in which courts apply intermediate or even strict scrutiny to generally applicable laws. A brief review of the cases confirms this thesis.

The Court in *Arcara v Cloud Books, Inc.*<sup>225</sup> attempted to explain why some incidental restraints, rather than falling outside the First Amendment's sphere of influence, receive the scrutiny usually given to content-neutral direct restrictions.<sup>226</sup> According to *Arcara*, generally applicable laws warrant First Amendment scrutiny in only two circumstances. First, such scrutiny is appropriate when the law, although not specifically referring to speech, "has the inevitable effect of singling out those engaged in expressive activity."<sup>227</sup> The Court cited as an example the imposition of a tax on newsprint, which would fall disproportionately—indeed, almost exclusively—on speakers. Second, review is proper when the conduct that draws the sanction has a significant expressive element. The Court used *O'Brien* as an example, noting that the act regulated in that case—the destruction of a draft card—itself "carr[ie]d] a message."<sup>228</sup> The Court cited in contrast

---

<sup>224</sup> For discussion of this relation in the context of religious liberty, see *Church of the Lukumi Babalu Aye, Inc. v City of Hialeah*, 508 US 520, 532-36 (1993).

<sup>225</sup> 478 US 697 (1986).

<sup>226</sup> Viewed solely in historic terms, this question has proved unimportant. Although the Court sometimes has subjected incidental restrictions to intermediate scrutiny, the Court never has overturned a restriction reviewed under this standard. See Schauer, 26 *Wm & Mary L Rev* at 787-88 (cited in note 207). When applied to incidental restraints, intermediate scrutiny has acquired a peculiarly toothless quality. Still, the decision to subject certain incidental restraints to heightened scrutiny has the potential to matter.

<sup>227</sup> *Arcara*, 478 US at 707.

<sup>228</sup> *Id.* at 702-03. Of course, destruction of a draft card does not always carry a mes-

cases in which the sanction, although affecting expressive activity, stemmed from nonexpressive conduct—as where a city closed a bookstore or fined a newspaper for violations of a zoning law or minimum wage ordinance.

The relevance of hugely disproportionate impact to the level of scrutiny is, under a motive-based approach, no great mystery. What separates direct from incidental restraints is breadth: whether the law applies to more than just speech. If an incidental restraint has no such sweep, effectively regulating only speech, then the danger it poses of illicit motive approaches the level associated with direct restraints, and the same standard of review thus should obtain.

A similar purpose-based analysis applies to the second factor the Court deemed important: whether the legal sanction results directly from, or merely impedes, expressive activity. The significance of this distinction is by no means transparent, nor does it fully explain the case law.<sup>229</sup> But if the distinction matters, it does so because it relates to motive—here, of administrative and judicial actors. When expressive activity triggers application of a law, the expression and the legal violation become ineluctably

---

sage; if it did, banning the act would constitute a direct, rather than incidental, restraint on expression. The point here is only that the conduct drawing a sanction in the particular case expresses a message. The Court in *Arcara* cited two other cases as similar to *O'Brien* in this respect, though only one in fact is so. In *Clark v Community for Creative Non-Violence*, 468 US 288 (1984), the Court considered the application of a ban on sleeping in Lafayette Park to demonstrators who sought to sleep in the park in order to call attention to the plight of the homeless. Here, the banned activity—sleeping—indeed had an expressive element. In *United States v Albertini*, 472 US 675 (1985), the Court considered the constitutionality of a statute barring a person from reentering a military base after being ordered not to do so, as applied to a person who wished to reenter a base in order to distribute leaflets. In this case, contrary to the Court's analysis, the conduct that drew the sanction—the reentry—was not itself expressive, although the policy regulating reentry interfered with the leafleter's ability to engage in expressive activity. Another explanation is thus necessary to explain the decision. See text accompanying notes 230-34. A more recent case, *Barnes v Glen Theatre, Inc.*, 501 US 560 (1991), conforms to the *Arcara* analysis. There, the Court considered whether a state could apply a public indecency statute to a nude-dancing establishment. Because expressive activity (nude dancing) drew the legal penalty, the Court (except for Justice Scalia) understood the case to raise a First Amendment question. See *id.* at 565-66 (plurality opinion); *id.* at 581 (Souter concurring); *id.* at 587 (White dissenting). But see *id.* at 576-78 (Scalia concurring).

<sup>229</sup> The Court has subjected to First Amendment review some restrictions that merely burden, rather than come down directly upon, expressive activity; *Albertini* presents an example. See note 228. Conversely, courts find First Amendment review unnecessary in some cases where expressive activity draws the sanction. I doubt, for example, that courts would engage in First Amendment review of a city's decision to apply a vandalism law to a person who draws swastikas on a synagogue wall. And I am sure courts would decline to apply First Amendment scrutiny to the conviction (under a statute prohibiting the destruction of property) of a political terrorist for blowing up the Statue of Liberty.

intertwined; this linkage is usually both less visible and less tangible when application of the law merely burdens expression. To the extent this entanglement occurs, the danger grows that prosecutors and fact finders alike will consider the nature of the expression in discharging their functions. The danger of discriminatory enforcement in such a case becomes roughly equivalent to that present when a speaker arguably violates a direct content-neutral restriction on speech, such as an antileafletting ordinance. Because the closeness of the link between sanction and expression affects the danger of improper motive, so too does it shift the appropriate constitutional standard.

In the end, however, the Court's decision to apply intermediate review to certain incidental restrictions may result not so much from use of the *Arcara* test as from a visceral sense that an illicit factor entered into a governmental decision—whether legislative, administrative, or judicial. Consider when the Court has treated an incidental restraint as raising a cognizable First Amendment question calling for intermediate scrutiny. First, in *O'Brien*, the law at issue, viewed in light of both legislative and extralegislative history, gave many indications, acknowledged (though trivialized) by the Court, of censorial motivation.<sup>230</sup> Next, in *United States v Albertini*<sup>231</sup> and *Clark v Community for Creative Non-Violence*,<sup>232</sup> administrative officials interfered with the expression of persons (like the speaker in *O'Brien*) engaged in protest against governmental policies.<sup>233</sup> Finally, in *Barnes v Glen Theatre, Inc.*, the state itself created suspicion by announcing that it would apply the allegedly general law to some, but not all expressive activities.<sup>234</sup> In short, in all of the cases in which the Court has tested an incidental restraint against an intermediate standard of review, there have been signs of impermissible motive.

---

<sup>230</sup> See *O'Brien*, 391 US at 385-86. See generally Dean Alfange, Jr., *Free Speech and Symbolic Conduct: The Draft-Card Burning Case*, 1968 S Ct Rev 1.

<sup>231</sup> 472 US 675 (1985).

<sup>232</sup> 468 US 288 (1984).

<sup>233</sup> In *Albertini*, the speaker wished to protest the nation's nuclear-arms policy; he previously had destroyed military documents. 472 US at 677-78. In *Community for Creative Non-Violence*, the speakers wished to protest the government's treatment of the homeless. 468 US at 289. An additional concern of improper motive in that case arose from the circumstances surrounding the adoption (not just the administration) of the facially general regulation. See *id* at 315-16 (Marshall dissenting).

<sup>234</sup> 501 US 560, 590 (1991) (White dissenting). The State asserted that its public indecency statute prohibited nude (erotic) dancing but did not prohibit nudity in theatrical or operatic productions.

In other cases, the Court has gone further, subjecting apparently incidental restraints to strict scrutiny—again, for reasons relating to motive. One set of these laws I have mentioned before: content-neutral laws whose application turns on the communicative effects of expression.<sup>235</sup> Many of these laws are generally applicable, in the sense I use the term; for example, a law barring breaches of the peace applies to both expressive and nonexpressive activities that disturb public order. I have suggested earlier why strict scrutiny applies in these circumstances: because of the way in which the content of ideas triggers the application of such a law, the ideological views of officials and members of the public likely will influence its enforcement. Asking whether the application of a facially general law turns on communicative impact thus serves as a way of deciding whether bias tainted the application.<sup>236</sup>

The Court also will apply strict scrutiny to a law of general application when either the asserted justification or the only rational justification for the law (or an application of it) relates to the communication of a message. *O'Brien* suggested this approach in asking whether the governmental interest asserted is related “to the suppression of free expression,” properly understood to mean whether the interest is related to the suppression of specific messages.<sup>237</sup> If the asserted interest or the only rational interest for an action is of this kind, then a court can assume that the official taking the action indeed considered the desirability of restricting certain messages. And when this is true, the probability is high that bias tainted the decision. This is not because all interests relating to the content of ideas are improper; to the contrary, many such interests reflect the cognizable harm an idea produces. But when officials self-consciously consider the merits of restricting ideas—whether because a law on its

---

<sup>235</sup> See Section III.B.2.

<sup>236</sup> At first glance, these cases may look no different from a case such as *Community for Creative Non-Violence*, 468 US at 288, where expressive activity also triggered the legal sanction. But in these cases, the content of expression, rather than the simple fact of expression, produces the sanction; the difference is between application of a breach-of-the-peace statute to very loud expression and application of the same statute to expression whose provocative content has stirred public ire.

<sup>237</sup> 391 US at 377. This interpretation follows as a simple matter of logic. The test *O'Brien* proposes for a law whose justification does *not* relate to the “suppression of free expression” is essentially the test applied to content-neutral regulations of speech. To merit a stricter standard of review, a law would need to have a justification relating not to the restriction of speech generally (which all content-neutral laws have), but to the restriction of speech of a certain content.

face effects this restriction or because the law, though written generally, has this restriction as its stated or obvious object—bias easily influences the evaluation of neutral harm-based criteria. An asserted or inescapable content-based interest, like content-based language, thus demands use of strict scrutiny.<sup>238</sup>

The point may become clearer by considering the legality of laws banning flag burning.<sup>239</sup> Note first, as illustration of a critical point of this Section, that the government may stop protesters from burning flags by enacting a general restriction—say, a ban on lighting fires in public places. But the government may not specifically proscribe the burning of flags for purposes of protest. The effects of these laws are of course no different. The dissimilar treatment arises, and must arise, from the disparate concerns about governmental motive raised by one law that sweeps broadly and another that focuses on expressive activity.

But now consider a harder case—a case involving the constitutionality of a law prohibiting any person from knowingly mutilating the flag of the United States.<sup>240</sup> The greater difficulty of the case again demonstrates the importance of motive, for the statute, when compared to the other two mentioned, has no different effect on expressive activity. What makes the case hard is that the statute falls near the line between incidental and direct restraints and thus raises a question about how far to suspect the motives of government. On one argument, the law applies on its face to an activity generally, regardless whether expressive in purpose or function; the law covers alike the person who burns a flag to protest a war and the person who uses flags for kindling.<sup>241</sup> But on the contrary argument, all of the rational inter-

---

<sup>238</sup> The same kind of analysis applies to a facially general law that applies to nothing, or almost nothing, but speech of a certain content. Just as a facially general law that operates to restrict only speech (and not conduct) ought to confront intermediate scrutiny, see text accompanying notes 227-29, so a facially general law that operates to restrict only speech of a particular kind ought to confront the strictest review. Arguably, the law in *O'Brien* was of this nature and ought to have been treated as if it exclusively related to antiwar protest, rather than encompassed as well "the odd soul who burns a draft card just to stay warm or to light up his campsite" or to deliver some other message. Tribe, 1993 S Ct Rev at 34 (cited in note 78). The same might be true of all the variations of flag-burning statutes I discuss below.

<sup>239</sup> The Court's two recent cases on the subject are *Eichman*, 496 US 310, and *Johnson*, 491 US 397.

<sup>240</sup> The hypothetical law provided here is a simplified version of the Flag Protection Act of 1989, 18 USC § 700 (1994), at issue in *Eichman*, 496 US at 314.

<sup>241</sup> This feature of the Flag Protection Act made *Eichman* more difficult than *Johnson*, in which the Court a year earlier had invalidated Texas's flag-burning statute. The Texas law prohibited flag burning only when it would cause offense to others—that is, only when it functioned as communication. See *Johnson*, 491 US at 400 n 1.

ests underlying the law relate to the restriction of a message.<sup>242</sup> In such a case, an indicator of illegitimate motive (content-based justification) undermines the indicator of legitimate motive (general application).<sup>243</sup> This is why the Court, in invalidating the statute, made the correct decision.

A brief discussion of another question recently decided by the Court—the constitutionality of hate-crimes laws—may serve as a summary of all these issues. Hate-crimes laws, as usually written, provide for the enhancement of penalties when a specified crime (say, assault) is committed because of the victim's race, religion, or other listed characteristic.<sup>244</sup> In *Wisconsin v Mitchell*,<sup>245</sup> the Court unanimously ruled that these laws present no First Amendment issue. The analysis so far helps to explain why the Court reached (and was right to reach) this decision.<sup>246</sup>

The key lies in the fact that the typical hate-crimes law—unlike a hate-speech law, as in *R.A.V.*—is an incidental restraint. On its face, the hate-crimes law targets not only speech, but a range of activity; it applies regardless whether the conduct at issue expresses a message. In this way, a hate-crimes law functions in the same way as any discrimination law—for example, in the sphere of employment relations.<sup>247</sup> When an

---

<sup>242</sup> See *Eichman*, 496 US at 315-16. Reasonable people can disagree—in fact, have disagreed—with this conclusion. See, for example, Geoffrey R. Stone, *Flag Burning and the Constitution*, 75 Iowa L Rev 111, 119 (1989) (proposing, with respect to a carefully drafted flag-burning statute, a reasonable interest unrelated to the communication of a particular message). I do not wish to rehash this debate, although I find the Court's position more persuasive. The critical points here concern the consequences of finding that the interests supporting a facially general law relate to the suppression of ideas, and the reason why those consequences follow.

<sup>243</sup> Again, the content-based justification is only an indicator of impermissible motive, not the thing itself. Many justifications relate to the suppression of messages, but are permissible because they relate to material harms the speech causes; consider, for example, a justification for a flag-burning statute that focuses on the tendency of such speech to provoke unlawful conduct. The problem of motive arises from the difficulty of evaluating this legitimate interest uninfluenced by ideological views of the speech in question. This problem will be evident in any case where the government's asserted interest in a law refers to a particular message, just as it is where the law itself makes this reference.

<sup>244</sup> See, for example, Cal Penal Code § 422.7 (West 1988 & Supp 1995); Wis Stat Ann § 939.645 (West Supp 1994).

<sup>245</sup> 508 US 476 (1993).

<sup>246</sup> For an excellent treatment of this question, see Tribe, 1993 S Ct Rev at 4-11 (cited in note 78). See also Kagan, 60 U Chi L Rev at 884-87 (cited in note 61).

<sup>247</sup> The Supreme Court in *Mitchell* recognized the analogy between Title VII and a hate-crimes statute. See *Mitchell*, 508 US at 487. It is noteworthy that both laws apply not only irrespective of whether the discrimination at issue expresses a message, but also irrespective of whether the discrimination stems from particular beliefs. If, for example, discrimination laws prohibited discharges or assaults motivated by racial hatred—rather than based on race—they would pose a more severe First Amendment problem.

employer fires an employee on the basis of race, the government may impose sanctions whether or not speech accompanies or itself effects the discharge; whatever speech is involved is incidental to the activity (race-based discharge) that the law condemns. The government may do the same when one person assaults another on the basis of race, again whether or not speech accompanies the conduct; a penalty enhancement may follow because it is pegged to conduct (race-based assault) that the state is attempting to prevent irrespective whether it has an expressive component. In both cases, the generality of the law provides a qualified assurance that disapproval of ideas qua ideas played no causal role in the legislative process.

One objection to this analysis might focus on the extent to which speech—more, speech of a certain kind—accompanies the prohibited conduct: if racist expression always or almost always is associated with race-based assault, then proscribing the activity amounts to proscribing the expression, with all the constitutional issues such a policy raises. In such a case, the so-called generally applicable law is not generally applicable after all; because the law applies only to certain expressive activity, the reasons for trusting the law disappear.<sup>248</sup> But I do not think this description accurately characterizes hate-crimes laws, which ban conduct that may and often does occur independent of expression; indeed, persons committing race-based crimes may try hard to conceal, rather than express, the racism inherent in the conduct.<sup>249</sup> Here, communication is neither so integral to nor so coincidental with the condemned activity as to reverse the usual presumption supporting generally applicable regulations.

A second objection to the analysis might point to the uneven way in which a hate-crimes law affects speech (when it does affect speech), effectively barring racist ideas and not others. There is no way to deny this skewing effect, and if it matters, then it calls for reversal of *Mitchell*. But I think it should not make such a difference. Many generally applicable laws affect speech in an asymmetrical way, as the conduct proscribed captures the expression of only certain messages.<sup>250</sup> Consider, for example, what kind of speech is likely to accompany a race-based discharge.<sup>251</sup> But if the law applies to conduct generally, the

---

<sup>248</sup> See text accompanying notes 227-29, 238.

<sup>249</sup> Tribe, 1993 S Ct Rev at 9-10 (cited in note 78).

<sup>250</sup> See text accompanying notes 219-24.

<sup>251</sup> Examples can proceed ad infinitum. Consider, to note two more, what kind of

critical barrier to the intrusion of illicit motive remains intact. That barrier is the focus of the law on acts irrespective of expression—a focus that usually prevents attitudes toward a message from influencing the legislative outcome. Again, then, the usual principles applicable to incidental restrictions seem to hold with respect to hate-crimes statutes.

The last objection to the analysis also is the strongest: that the only rational justification for a hate-crimes law relates to the message the proscribed activity conveys. The Court rightly saw flag-burning laws in this light—as an effort, underneath the cover of an incidental restraint, to suppress communication of a message.<sup>252</sup> Perhaps the same argument applies to hate-crimes laws; indeed, the Court in *Mitchell*, though upholding such a law, understood it in much this way, pointing to interests the government had in restricting expression of racist messages.<sup>253</sup> But this view of hate-crimes laws is not necessary. The government may have a non-speech-related interest for sanctioning race-based assault, no less than race-based discharge: an interest in eradicating racially based forms of disadvantage—in preventing disproportionate harm from falling, by virtue of status alone, on members of a racial group. Given this interest, existing apart from any speech, the Court correctly treated hate-crimes laws as laws of general application.

---

speech falls victim to a law banning treason or prohibiting the assassination of the President.

<sup>252</sup> For an example whose structure parallels a hate-crimes law, consider a penalty enhancement provision applicable to persons who obstruct voting on the basis of a voter's membership in the Republican Party. In this case too, the government's interest in the law cannot but relate to favoring or disfavoring particular viewpoints. But now consider a broader law applying enhanced penalties to persons who obstruct voting based on the voter's affiliation with *any* political party. Such a law could have emerged from a governmental interest in protecting persons from suffering disproportionate harm as a result of their political views, analogous to the interest I will soon note in protecting persons from suffering disproportionate harm as a result of their race. Accordingly, this law would meet constitutional standards: it applies regardless whether the conduct communicates a message, and the government has a credible interest in the law not related to favoring or disfavoring particular messages.

<sup>253</sup> The Court noted that race-based crimes were more likely than other crimes "to provoke retaliatory crimes, inflict distinct emotional harms on their victims, and incite community unrest." See 508 US at 488. But these effects arise, largely if not entirely, from the expressive component of a race-based crime—the aspect of the activity that communicates, and is meant to communicate, a message. The interest in preventing these effects is of course not illegitimate. But it is an interest related to restricting a message. Thus, if this interest alone lies behind hate-crimes laws, then courts, however they rule on the legitimacy of the laws, should regard them as restrictions of expression.

But this analysis of hate-crimes legislation is no more than an illustration. The basic points I wish to make concern not whether courts ought to regard any particular piece of legislation, including hate-crimes laws, as falling within the category of incidental restraints, but instead what rationale lies behind that category and what limits that rationale sets on the category's boundaries. The doctrine of incidental restraints, like so much of First Amendment law, arises from a desire to flush out impermissible purposes on the part of the government. The reason why incidental restraints usually receive no First Amendment review relates to the low risk that hostility or partiality toward ideas tainted these restrictions. The circumstances in which incidental restraints confront a measure of scrutiny—and the level of scrutiny then employed—also connect to the risk of taint in certain laws. Once again, the doctrine acts as a complex mechanism to provide review where necessary, and of the kind necessary, to invalidate improperly motivated governmental actions.

#### IV. THE UNDERPINNINGS OF MOTIVE ANALYSIS

The primary task of this Article is to lay out a descriptive theory of the First Amendment—an explanation of the state of First Amendment doctrine, an account of the wellspring of the law. I cannot here proceed much beyond the aims of this project to resolve all of the normative questions this theory raises. In particular, I cannot here provide a full justification for structuring First Amendment law around the question of motive, rather than around the question of effects, whether on a speaker or an audience. But I also cannot conclude this Article without considering briefly what might lie behind the law's focus on motive—why, that is, motive might (or, at least, might be thought to) matter. In this Section, I turn to the normative underpinnings of motive analysis.<sup>254</sup>

The importance of motive in First Amendment analysis, as in other spheres of constitutional analysis, is in many ways mysteri-

---

<sup>254</sup> Many scholars have discussed in detail the appropriate role of governmental motive in constitutional law, especially in cases involving the Equal Protection Clause. See, for example, Lawrence A. Alexander, *Introduction: Motivation and Constitutionality*, 15 San Diego L Rev 925 (1978); Brest, 1971 S Ct Rev at 95 (cited in note 77); Theodore Eisenberg, *Disproportionate Impact and Illicit Motive: Theories of Constitutional Adjudication*, 52 NYU L Rev 36 (1977); Ely, 79 Yale L J at 1205 (cited in note 40); Michael J. Perry, *The Disproportionate Impact Theory of Racial Discrimination*, 125 U Pa L Rev 540 (1977). For a brief but excellent discussion, focusing on the role of motive in First Amendment analysis, see Tribe, 1993 S Ct Rev 1 (cited in note 78).

ous. Consider two laws with identical real-world consequences, either as to the aggregate quantity of speech or as to the quality of the speech market. An analysis focusing on motive may force us to treat these laws differently, even though they affect us in the same manner. Or consider two laws with diverse impacts, one restricting the sum total of speech more than the other or distorting the sphere of discourse in more dramatic fashion. An analysis focusing on motive may force us to treat these laws identically, even though we experience the one as more confining than the other. To say this much is only to acknowledge the inevitable outgrowth of motive analysis—only to recognize the potential disjunction between the reasons for doing a thing and the results of having done it. The question underlying motive-based analysis is why we should focus in this way on what led to an action, not what the action accomplished.<sup>255</sup>

One way to approach this issue is to consider the kinds of justifications we usually offer for rules of process. (A prohibition relating to motive is, after all, nothing more than such a rule, operating to exclude certain factors from the decision-making process.) First, we may adopt such a rule because it will promote good consequences, where the criterion for deciding what constitutes a good consequence comes from outside the rule itself. Second, and alternatively, we may adopt a rule of process for its own sake, because it possesses certain attributes or expresses certain norms, the correctness of which renders any outcome it produces correct as well. The distinction is between rules of process whose justification derives from the appropriateness of the results they promote, as independently defined, and rules of process whose justification derives from internal attributes, which themselves define what results count as appropriate.<sup>256</sup>

---

<sup>255</sup> The famous Holmesian quip respecting the importance of motive tends to glide over this problem. "[E]ven a dog," Holmes wrote, "distinguishes between being stumbled over and being kicked." Oliver Wendell Holmes, *The Common Law* 7 (Belknap 1963) (Mark DeWolfe Howe, ed). But the injured dog might think that this distinction, however easy to make, ought not to carry weight. It is one thing to distinguish and another to make the distinction matter in a legal system. The critical question is why this distinction should matter given that the resulting bruise seems the same. Holmes's epigram is perhaps suggestive of an answer, as I later discuss, but does not itself provide it.

<sup>256</sup> The discussion in the preceding paragraph follows John Rawls's typology of systems of procedural justice. Rawls defines "imperfect procedural justice" as a situation in which procedures gain their correctness from their ability to promote certain independently justified outcomes. He defines "pure procedural justice" as a situation in which procedures lend their own (internal) correctness to the outcomes they produce. See John Rawls, *A Theory of Justice* 85-86 (Belknap 1971). This distinction correlates roughly to one between consequential and deontological justifications for a rule—that is, justifications

Using this framework, two kinds of theories may explain the prevalence in First Amendment law of motive-based analysis. First, courts may focus on motives because doing so will promote good outcomes, as defined by some independent set of criteria. On this theory, the gulf I have presupposed between the reasons for an action and the effects of an action in an important sense closes. The reason to think about reasons has to do with the likelihood that the consideration of certain reasons will systematically and predictably lead to actions that have adverse consequences.<sup>257</sup> Second, courts may focus on motives because motives, in and of themselves, are what matter—so much that any actions deriving from improper motives also become improper, by a kind of automatic motion. On this theory, the division I have posited between reasons and effects closes in a different but no less fateful manner. Now, the only effects that matter are the effects (whatever they may be) that emerge from particular reasons for action.<sup>258</sup> Both of these theories are as yet mere outlines; the remaining questions concern how to give them content.<sup>259</sup>

A consequentialist theory of motive analysis must provide an account of how certain motives foster adverse outcomes, as defined by an independent criterion of value. Assume here that we should assess outcomes in the fashion of the audience-based model: the optimal state of public discourse is that most illuminating to and desired by an ideally curious and engaged audience. Now consider how the concept of improper motive I have developed relates to this view of desirable (and undesirable) outcomes. When self-interest or ideological hostility enters into a restriction on speech, the odds increase that the resulting action will impoverish the sphere of public discourse.<sup>260</sup> By happen-

---

that refer to the desirability of producing some independently justified value and justifications that operate autonomously of any particular set of consequences.

<sup>257</sup> This closure, I hasten to add, does not undermine the thesis of this Article. As I explain, the distinction between motive-based analysis and effects-based analysis remains all-important for purposes of constructing (and explaining) First Amendment doctrine.

<sup>258</sup> The apparent oddity that good motive can save a law with detrimental effects (and bad motive doom a law with beneficial effects) thus disappears. On this account, in its purest form, the presence or absence of impermissible motive itself defines whether effects are detrimental or beneficial.

<sup>259</sup> These two accounts can coexist, if not in their purest forms, then in some modified versions. In a combined approach, the insistence on a rule of process would derive from both the value inherent in the rule and the connection of the rule to certain outcomes.

<sup>260</sup> See text accompanying notes 100-01. The argument here is that improperly motivated action will tend to distort the speech market, not that content-based action will tend to distort the speech market. I have criticized the latter argument many times, noting that content-based regulation sans bad motive (say, if conducted on a random

stance, some improperly motivated restrictions will enhance, and some purely motivated restrictions will mutilate, the thinking process of the community. But in general, a system in which the government freely may restrict ideas on the ground that they challenge the power or wisdom of officials will produce a less healthy debate than a system in which the government has no such ability. A rule proscribing actions arising from censorial motive thus will promote the set of outcomes that the audience-based model deems desirable.

But even assuming this claim is accepted, the reason to focus on reasons is still not self-evident. If the ultimate question relates to the effects of an action on public discourse, then courts seemingly should make just this inquiry. Asking about improper motive so as to hazard a guess about untoward effects appears a strangely circuitous way of addressing the issue. For the consequentialist theory under discussion to work, an inquiry into audience-based effects must be infeasible; only then might discovering reasons provide the most judicially manageable way to evaluate an action's impact on public discourse.

An argument of this kind is indeed plausible, given the nature of the inquiry into effects and the limits of judicial capacity. The criteria that the audience-based model provides by which to judge the results of a speech regulation may be insufficiently definite and detailed to lend themselves to direct application. Consider that we do not possess a fully developed sense of what an optimal marketplace of ideas would look like.<sup>261</sup> We have instead a set of hazy generalities (rich, robust, balanced, diverse) by which we refer to the ideal state. These standards, although sometimes of concrete assistance, often cannot tell us whether a governmental action, the very operation of which may be uncertain, will impair or improve public discourse. And even if they always could do so in the hands of, say, Hercules, they seem far too diffuse—far too manipulable—to trust to a judge with her own set of interests and biases. The problem with an effects-based standard is one of judicial administration. The questions it forces judges to ask about what ideas are over- or underrepre-

---

basis) is as likely to improve as impair the speech market. All I am claiming here is that when the government restricts ideas because officials see them as threatening or distasteful, the action more often will debilitate than strengthen public discussion.

<sup>261</sup> We also may not possess precise ways to measure or describe the current distribution of ideas, so that we could determine the difference between it and the ideal. For fuller discussion of these issues, see Strauss, 1993 U Chi Legal F at 202-10 (cited in note 20).

sented, about who has talked too much or too little, about when "drowning out" has occurred, are not subject to unbiased, reliable evaluation.

If this is so, the focus on governmental motive that marks First Amendment doctrine may function as a kind of proxy for an inquiry into the effects of a restriction of speech on an audience. This does not mean that motive-based analysis plays a subsidiary role to effects-based analysis or that the two meld into each other.<sup>262</sup> To the contrary, it is the search for improper motive that drives the doctrine. That search generates a set of doctrinal rules different from the rules that would flow from a direct effects-based inquiry. Those rules then spawn a set of results different from the results that would emerge from effects-based doctrines. Yet beneath the operation of these rules indeed may lie a concern with consequences. The focus on motive, on this account, provides an indirect way of identifying actions with untoward effects on public discourse. This identification mechanism is necessarily imprecise—both over- and underinclusive. But given the difficulties of inquiring directly into effects, it may be the best such instrument that courts can find.

The notion here should seem familiar, for I have based most of this Article on a similar form of argument. What I have said already goes something as follows: We wish to discover improper motive, but cannot do so by making a direct inquiry; we instead construct a set of rules, turning on the facial terms of legislation, to identify motive indirectly; we realize these rules will prove imprecise, capturing too much and too little, but we use them because we can think of no better way to discover improper motive. What I am positing now adds an anterior, but similarly constructed, line of reasoning to this one. Why do we wish to discover improper motive? Perhaps because we wish to discover adverse effects, but cannot do so directly; because we know that actions tainted with certain motives tend to have such consequences; because although a focus on motive will prove imprecise, we can think of no better way to gauge the effects of an action on the state of public discourse. Hence we emerge with a set of First Amendment rules serving as "double proxies"—first, and more

---

<sup>262</sup> As Larry Alexander has noted, there is a difference between motive theories and effects theories even when, as is often the case, the proscribed motives are "selected because of the social effects associated" with them. Alexander, 15 San Diego L Rev at 931 (cited in note 254). The critical aspect of motive theories is that "in the final analysis, where motives and effects are inconsistent, the motives, not the effects, govern." *Id.*

proximately, for an inquiry into a certain kind of motive; then, and more remotely, for an inquiry into a certain kind of effect. At each step, something gets lost; the rules function imperfectly to flush out improper motive, still more imperfectly to identify adverse effects. But on this theory, we use the rules because they, better than any others, allow the discovery of improper motive—in order to achieve the discovery of adverse impact.

In contrast to this consequentialist account, the second explanation for the law's focus on motive is purely internal. Here, the reasons underlying a law do not provide evidence of something else, but themselves constitute the ultimate issue. But this raises the question why motives ought so to matter—why motives, for their own sake and irrespective of material consequence, should determine the legitimacy of governmental action.

The answer may begin to emerge if we recognize that two actions having similar material outcomes may express different values and have different meanings.<sup>263</sup> This contrast, I think, is what Holmes meant to highlight when he distinguished between stumbling over and kicking a dog.<sup>264</sup> The former may suggest a lack of optimal care, but the latter suggests contempt or hatred.<sup>265</sup> The same dynamic between reasons, actions, and meanings often arises. Consider the difference between a policy that intentionally excludes African-Americans from employment and a test that as effectively, but unintentionally, prevents them from gaining desired positions. There are, of course, (effects-based) reasons to treat these actions identically; but it is wholly intelligible to say that the former is worse than the latter because it conveys an attitude of disrespect or malevolence. (Indeed, equal protection law, for this reason, treats deliberate discrimination

---

<sup>263</sup> See Elizabeth Anderson, *Value in Ethics and Economics* 33-34 (Harvard 1993) (“[T]he distinction[ ] . . . between foreseeing and intending certain consequences . . . mark[s] [a] distinction[ ] in the expressive significance of actions.”); Joseph Raz, *The Morality of Freedom* 378 (Clarendon 1986) (Certain acts have “meaning regardless of their actual consequences . . . expressing disregard or even contempt.”); Richard H. Pildes, *Avoiding Balancing: The Role of Exclusionary Reasons in Constitutional Law*, 45 *Hastings L J* 711, 728 (1994) (“How an action comes about shapes what it means and therefore what it is.”).

<sup>264</sup> See note 255.

<sup>265</sup> The dog also may fear the kick more than the stumble for related, but entirely consequentialist reasons. Even if the bruise from the two is the same, the contempt inherent in the kick increases the likelihood that yet another kick will come and yet another bruise follow. This reasoning suggests a connection between First Amendment law's focus on motive and its amenability to slippery slope arguments. We care about bad reasons because they entail repeated bad results; we fear the occurrence of repeated bad results because we suspect the existence of bad reasons.

having trivial consequences as more problematic than incidental discrimination having great impact.<sup>266</sup>) An action acquires meaning in part through motive, and the meaning of an action in part defines it. Hence, what an action *is* derives not only from what it does, but from where it comes from; so too, then, whether the action is legitimate involves this matter.

The doctrine of impermissible motive, viewed in this light, holds that the government may not signify disrespect for certain ideas and respect for others through burdens on expression. This does not mean that the government may never subject particular ideas to disadvantage. The government indeed may do so, if acting upon neutral, harm-based reasons. But the government may not treat differently two ideas causing identical harms on the ground that—thereby conveying the view that—one is less worthy, less valuable, less entitled to a hearing than the other. To take such action—in effect, to violate a norm of ideological equality—would be to load the restriction of speech with a meaning that transcends the restriction's material consequence.

The First Amendment's focus on motive, on this account, serves as an analogue in the speech context to the principle that the government must treat all persons with equal respect and concern.<sup>267</sup> This principle, which may well explain much of equal protection law,<sup>268</sup> holds in part that the government may not treat some persons differently from others because they are deemed less intrinsically worthy. If such impermissible considerations intrude into the decision-making process, the results of that process likewise become improper; this is so even if the same results could have stood had hostility not infected the process.<sup>269</sup> So too here, except that the principle of impartiality applies not to persons, but to ideas. In determining whether to restrict speech, the government may not rank the worth or "rightness" of messages; to do so would be to register a kind of disrespect that automatically renders the action improper.

---

<sup>266</sup> See, for example, *Washington v Davis*, 426 US 229, 244-48 (1976).

<sup>267</sup> See, for example, Ronald Dworkin, *Taking Rights Seriously* 272-73 (Harvard 1977) (proposing that government may not treat persons unequally "on the ground that some citizens are entitled to more because they are worthy of more concern").

<sup>268</sup> See Cass R. Sunstein, *Public Values, Private Interests, and the Equal Protection Clause*, 1982 S Ct Rev 127, 143.

<sup>269</sup> We might think of officials in such a case as having violated the rules of a game. It is no excuse to a charge of cheating to say that the same results would be unobjectionable in a game played honestly. The fact of cheating renders the result improper, whatever the result might be. See Rawls, *A Theory of Justice* at 86 (cited in note 256).

But this account has left open a crucial issue: why is it improper for the government, through restrictions on speech, to show contempt for contemptible ideas, independent of the harm they cause? One answer to this question negates its premise: this answer insists that in life, as in law, "there is no such thing as a false idea," neither an abhorrent one;<sup>270</sup> or, to put the point somewhat less baldly (in the way Justice Holmes flirted with it), the very, and only, definition of truth and wisdom is what emerges from free discussion.<sup>271</sup> But this answer entails an extreme skepticism, unacceptable to most of us because incompatible with a host of our considered judgments. It explains the First Amendment principle of equality only by assuming a world of moral indeterminacy (thankfully) impossible to recognize. A second and better answer to the question refers to the probability that the government will err, as a result of self-interest or bias, in separating the true and noble ideas from the false, abhorrent ones; a scheme of neutrality thus provides the surer means to make this distinction. But this answer largely returns us to the consequentialist basis for focusing on motive; again, what is stressed is the connection between distrusting government and achieving the best possible public discussion. What I am trying now to explore is a different rationale for the focus on motive.

Two versions of this rationale seem possible, both referring to the locus of decision-making authority in our political system, but one sounding in terms of individual rights, the other in terms of popular sovereignty.<sup>272</sup> In the first, the prohibition of ideological motive, and the correlative duty of equal respect for ideas and their proponents, serves as just one application of a general ban on subjecting people to disadvantage for reasons that do not relate to harm, but instead arise from judgments of moral value (or from official self-interest). The government, in the view of many liberal theorists, cannot disadvantage a person because the way she lives is immoral or repellant, even if it is so—or because others view it as immoral or repellant, even if they do.<sup>273</sup> (Nor

---

<sup>270</sup> *Gertz v Robert Welch, Inc.*, 418 US 323, 339 (1974).

<sup>271</sup> See *Abrams v United States*, 250 US 616, 630 (1919) (Holmes dissenting); *Gitlow*, 268 US at 672-73 (Holmes dissenting).

<sup>272</sup> Richard Pildes has suggested that motive inquiries throughout constitutional law serve "to set the boundaries between separate spheres of authority"—to aid in the "differentiation of political authority that is crucial to liberalism and the Constitution." See Pildes, 45 *Hastings L J* at 713, 715 (cited in note 263). Both of the following accounts can be understood in this light.

<sup>273</sup> See John Stuart Mill, *On Liberty*, in J.M. Robson, ed, 18 *Collected Works: Essays on Politics and Society* 213, 223-24 (Toronto 1977); Ronald Dworkin, *A Matter of Principle*

can the government impose such a disadvantage, in the absence of any harm, merely because public officials thereby would gain.) If this is so, it follows that the government cannot disadvantage a person because what she thinks or says is immoral or repellant or because others view it as such.<sup>274</sup> The narrower (speech-related) principle inheres in the broader; both are aspects, so the argument goes, of the appropriate relationship between the government and individuals within a liberal society.<sup>275</sup>

The second kind of nonconsequentialist account for the prohibition of ideological motive relates more exclusively to expression, emphasizing the place of such activity in a democracy.<sup>276</sup> On this view, the prohibition of ideological motive, and its concomitant principle of equality, lies at the core of the First Amendment because it lies at the core of democratic self-government. The democratic project is one of constant collective self-determination; expressive activity is the vehicle through which a sovereign citizenry engages in this process by mediating diverse views on the appropriate nature of the community. Were the government to limit speech based on its sense of which ideas have merit, it would expropriate an authority not intended for it and negate a critical aspect of self-government.<sup>277</sup> Democracy demands that sovereign citizens, through each generation, retain authority to evaluate competing visions and their adherents—to decide which ideas and officials merit approval. Hence democracy bars the

---

191, 203 (Harvard 1985); Raz, *The Morality of Freedom* at 420 (cited in note 263).

<sup>274</sup> For the most persuasive statement of this position, see Dworkin, *A Matter of Principle* at 353 (cited in note 273) ("People have the right not to suffer disadvantage . . . just on the ground that their officials or fellow-citizens think that their opinions about the right way for them to lead their own lives are ignoble or wrong.").

<sup>275</sup> Joseph Raz has argued that a liberal government may make moral judgments in the absence of harm by means other than coercion. In his words, the harm principle holds that "[w]hile [moral] ideals may indeed be pursued by political means, they may not be pursued by the use of coercion except when its use is called for to prevent harm." See *The Morality of Freedom* at 420 (cited in note 263). In the First Amendment context, such an argument might support a decision to apply purpose analysis more strictly when the government restricts speech than when the government funds speech or speaks itself.

<sup>276</sup> The thoughts in this paragraph owe much to the work of Robert Post and Geoffrey Stone. See Robert Post, *Managing Deliberation: The Quandary of Democratic Dialogue*, 103 *Ethics* 654, 660-61 (1993); Robert C. Post, *Racist Speech, Democracy, and the First Amendment*, 32 *Wm & Mary L Rev* 267, 290-91 (1991); Stone, 131 *Proc Am Phil Soc'y* at 253 (cited in note 137).

<sup>277</sup> This is true even when the government is acting as representative of the majority of citizens. The notion of democracy used here means something different from simple majoritarian decision making. The notion invokes a continuing, evolving process by which sovereign citizens accommodate and reaccommodate diverse views and, in so doing, create and recreate their society. See Post, 32 *Wm & Mary L Rev* at 279-83 (cited in note 276); Stone, 131 *Proc Am Phil Soc'y* at 253 (cited in note 137).

government from restricting speech (as it also bars the government from limiting the franchise) on the ground that such activity will challenge reigning beliefs or incumbent officials. The government must treat all ideas as contingent, because subject to never-ending popular scrutiny. On this view, the prohibition of certain motives again serves as a way to delineate the proper sphere of authority, hereby preventing a democratic state from contravening key principles of self-government and thereby undermining its foundation.<sup>278</sup>

All of these explanations go to the question why governmental motive is important. A separate and harder question concerns why motive should be all-important, such that an inquiry into motive precludes an inquiry into effects. I have considered this question briefly in discussing audience-based effects, noting that an inquiry into motive tests these effects better, though more circuitously, than a frontal inquiry could do. I have neglected this question altogether in discussing speaker-based effects.<sup>279</sup> But for two reasons, I will not attempt here to address this question further. First, I have never claimed that current law wholly excludes an inquiry into effects; I have claimed that motive plays the dominant, not the exclusive, role in the doctrine. Second, I have never proposed to show that the most sensible system of free expression would focus on issues of governmental motive to the extent our system does, let alone to the exclusion of all others. I have posited only that our system of free expression focuses on motive—and, to buttress that claim, noted the normative commitments that underlie this election. I leave for another day the question whether our doctrine, in attempting to discover improper motive, has neglected too much else of importance.

#### CONCLUSION

In his opinion for the Court in *R.A.V.*, Justice Scalia distinguished the case before the Court from several of his own inven-

---

<sup>278</sup> An argument of this kind also might support applying motive analysis differently to governmental efforts to restrict private speech and governmental efforts to participate in the speech market. See note 275. It is arguable (though subject to many limitations and exceptions) that the government does not expropriate a sovereign citizenry's ultimate authority to decide which ideas are worthy when the government acts not to limit a debate, but to engage in it.

<sup>279</sup> One explanation for relegating speaker-based effects to a secondary position is that we can and do expect the political process, in the absence of impermissible motive, to protect against inordinate restrictions on expressive opportunities.

tion.<sup>280</sup> If the case had involved secondary effects, Justice Scalia noted, a different issue would have been presented. So too if the case had involved an incidental restriction. In these and other circumstances, Justice Scalia wrote, the Court would have countenanced governmental action burdening only certain messages. The concurring opinions in the case castigated Justice Scalia's list of distinctions and exceptions. Justice Stevens wrote that the Court had "offer[ed] some ad hoc limitations" on its holding in order to contain the "perversities" it would engender.<sup>281</sup> Justice White similarly noted the Court's effort to "patch[ ] up its argument with an apparently nonexhaustive list of ad hoc exceptions."<sup>282</sup> But Justice Scalia's opinion did contain a rationale for his catalogue of distinctions. What he was trying to separate from the St. Paul ordinance, the Justice intimated, were laws that, although imposing differential burdens on ideas, "refute[d] the proposition" that they were "even arguably 'conditioned upon the sovereign's agreement with what a speaker may intend to say.'"<sup>283</sup> In essence, Justice Scalia was pointing toward a set of rules and categories (whether he got them exactly right is not what matters) that attempt to sort out, even if implicitly, regulations based solely on neutral determinations of harm from regulations tainted with ideological motive.

The reaction of Justice White and Justice Stevens is not surprising—nor is the more general sense of rebellion against the increasingly technical, complex classificatory schemes of First Amendment law, which Justice Scalia's opinion highlighted. One commentator wrote more than a decade ago of the "elaborate codification of the First Amendment," warning against "excess categorization."<sup>284</sup> Since then, the doctrine has become only more intricate, as categories have multiplied, distinctions grown increasingly fine, and exceptions flourished and become categories of their own. Little wonder that Justice Stevens could refer to the *R.A.V.* Court's "adventure[s] in . . . doctrinal wonderland."<sup>285</sup> Or that a recent commentator on Supreme Court jurisprudence, including *R.A.V.* and other First Amendment cases,

---

<sup>280</sup> 505 US at 388-90.

<sup>281</sup> *Id.* at 423 (Stevens concurring).

<sup>282</sup> *Id.* at 407 (White concurring).

<sup>283</sup> *R.A.V.*, 505 US at 390, quoting *Metromedia, Inc. v. City of San Diego*, 453 US 490, 555 (1981) (Stevens dissenting in part).

<sup>284</sup> Frederick Schauer, *Codifying the First Amendment: New York v. Ferber*, 1982 S Ct Rev 285, 288, 316.

<sup>285</sup> 505 US at 418 (Stevens concurring).

could criticize the Court for “devolv[ing] into conceptualism and technicality” and exhibiting “an almost medieval earnestness about classification and categorization.”<sup>286</sup>

But before we dismiss this conceptual scheme, we should at least explore what lies behind it. This means more than engaging in the now familiar debate about the relative merits of rules and standards. It means more than asking in the abstract whether formalism might have its uses. It means, instead, examining in concrete terms how these rules function and what they accomplish. That is what this Article has tried to do; for only when we know why the doctrine has emerged and what purposes it serves will we know whether and how to modify it.

What I have argued in this Article is that most of First Amendment doctrine constitutes a highly, but necessarily, complex scheme for ascertaining the governmental purposes underlying regulations of speech. The Court could not—and knows it could not—discover these motives through direct inquiry; in all but the most unusual case, the government could offer a permissible reason for its action, and the Court could not tell whether this reason was real or pretextual. Hence, the Court (whether consciously or not is unimportant) has constructed and relied upon a set of rules and categories, focusing on the facial aspects of a law, that operates as a proxy for this direct inquiry. Because these rules operate at a step removed, they are both over- and underinclusive. But they do well, if not perfectly, what could not be done in their absence—ferreting out and then invalidating impermissibly motivated governmental actions. The categories, the distinctions, and the rules of First Amendment law thus have a rationale and purpose not immediately apparent; if courts could evaluate motive directly, they could remove the lion’s share of the First Amendment’s doctrinal clutter.

The presence of this underlying principle, explaining and rationalizing First Amendment doctrine, does not make the doctrine self-evidently correct. We may believe that the doctrine cares too much about motive or that it cares too little about other things. But the principle does make the doctrine internally consistent and coherent. If the current doctrinal formulations are wrong, they are largely wrong as a whole and for the same reasons. And those reasons would relate—as most of the law re-

---

<sup>286</sup> Morton J. Horwitz, *The Supreme Court, 1992 Term—Foreword: The Constitution of Change: Legal Fundamentality Without Fundamentalism*, 107 Harv L Rev 30, 98 (1993).

lates—to the decision to treat the question of governmental motive as the preeminent inquiry of the First Amendment.

# When A Speech Code Is A Speech Code: The Stanford Policy and the Theory of Incidental Restraints

Elena Kagan\*

The title of Professor Grey's article, *How to Write a Speech Code Without Really Trying*, is instructive, if in some tension with what follows it. The title suggests two points: first, that Grey did not intend to write a speech code; second, that Grey wrote a speech code. I'll trust Grey on the first; he would know better than I. I'll agree with him on the second — except that I'm agreeing with his title only; as the rest of his article makes clear, Grey still denies he wrote a speech code. It is on that essential point, involving the distinction in First Amendment doctrine between direct and incidental restraints, that I take issue with his exceptionally interesting and provocative article.

Grey wrote an exceedingly narrow speech code — perhaps the narrowest that can be imagined. He wrote a speech code, as he insists, that in some sense recognized the value of a free speech system. He wrote a speech code that a reasonable system of First Amendment law could permit.<sup>1</sup> But Grey did write a

---

\* Professor of Law, University of Chicago; on leave 1995-96 to serve as Associate Counsel to the President. A.B. 1981, Princeton University; M. Phil. 1983, Oxford University; J.D. 1986, Harvard Law School. Thanks to Geoffrey Stone and David Strauss for helpful comments.

<sup>1</sup> This is not to say that the current system of First Amendment law permits the Stanford Policy. That Policy, as Grey explains, barred a subset of unprotected speech — specifically, fighting words, based on sex, race, or other listed characteristics. As restrictions on speech go, this one is narrow indeed; too, it is prefaced, for whatever this is worth, with a statement of commitment to the principles of free inquiry and speech. But unless Grey is right that the Stanford Policy should be viewed not as a ban on speech, but as part of a generally applicable regulation against discrimination, the Policy falls within the holding of *R.A.V. v. City of St. Paul*, 505 U.S. 377 (1992), that a prohibition of race-based fighting words violates the First Amendment. I have discussed that decision in an earlier article. See Elena Kagan, *The Changing Faces of First Amendment Neutrality: R.A.V. v. St. Paul, Rust v. Sullivan, and the Problem of Content-Based Underinclusion*, 1993 S. CT. REV. 29, 60-76. As I noted there, I agree with Grey and all the concurring Justices in *R.A.V.* that even under its

speech code, and from that fact a great deal both does and should follow.

This Comment on Grey's article addresses the scope of the First Amendment's doctrine of incidental restraints, which I think Grey misdescribes. It considers both the rationale and the need for that doctrine, which I think Grey underacknowledges. And finally it notes some practical political effects of the doctrine, which I wish Grey, in his capacity as drafter of the Stanford Policy, had more fully recognized. What is perhaps most disturbing about the Stanford experience is not that the University adopted, yes, a speech code, but that in doing so, it did little to foster, and perhaps much to undermine, its own (and Grey's own) goal of equality.

#### I. APPLYING THE DOCTRINE OF INCIDENTAL RESTRAINTS

Grey defends the Stanford Policy primarily on the basis of the distinction prevalent in First Amendment law between direct and incidental restraints on expression.<sup>2</sup> The Policy, according to Grey, did not concern speech as such; it concerned all discriminatory harassment, of which "hate speech," narrowly defined, formed just a part.<sup>3</sup> Because the Policy was generally applicable

---

own analysis, the *R.A.V.* Court might well have upheld the St. Paul ordinance — and thus also approved the Stanford Policy — as a ban on the subcategory of fighting words that most pose the dangers associated with fighting words generally.

<sup>2</sup> Grey's need to defend the constitutionality of the Policy arises from the Leonard Law, which applies First Amendment requirements to the disciplinary regulations of California's private universities. See CAL. EDUC. CODE § 94367 (West Supp. 1996). Even before passage of the Leonard Law, however, both Stanford and Grey had committed themselves to abiding by First Amendment standards. Whether a university like Stanford should commit itself in this manner seems to me a difficult question, which this Comment will not address.

<sup>3</sup> See Thomas C. Grey, *How to Write A Speech Code Without Really Trying: Reflections on the Stanford Experience*, 29 U.C. DAVIS L. REV. 891, 928-35 (1996). Grey assumes in his article, as I do in this reply, that an inarguably general law against discriminatory harassment — a law that did not mention speech at all — would meet any applicable First Amendment requirements, even when applied to such speech as the Stanford Policy covered. The Supreme Court has indicated its agreement. See *Harris v. Forklift Sys., Inc.*, 114 S. Ct. 367, 370 (1993). Some commentators, however, have disputed the point. See, e.g., Kingsley R. Browne, *Title VII as Censorship: Hostile-Environment Harassment and the First Amendment*, 52 OHIO ST. L.J. 481 (1991) (stating that broad judicial definition of harassment in Title VII, including speech, is inconsistent with First Amendment); Eugene Volokh, *Freedom of Speech and Workplace Harassment*, 39 UCLA L. REV. 1791 (1992) (arguing that general anti-harassment laws do not satisfy First Amendment requirements).

in this ma  
no seriou  
specifically  
that fighti  
sexual ori  
broader c  
should ha  
by making  
covered, t  
the Policy

To evalu  
and ask w  
and incide  
no sense  
the effects  
Stanford s  
will "suffer  
than from  
that cover  
conduct in  
what matte  
hear and  
about race  
from the  
the other  
the distinc  
Court's us  
the effects  
audience.<sup>6</sup>

---

<sup>4</sup> See Grey,

<sup>5</sup> For mor

*Purpose: The Ro*  
491-505 (1996)  
*Restraints on Co*  
*Neutral Restrictio*

<sup>6</sup> To use a  
Sunday edition  
fits for compan  
law is nil and th  
nications indust  
mere rationality

in this manner, applying to both speech and conduct, it raised no serious First Amendment problem. Of course, the Policy specifically described its application to expression, explaining that fighting words based on sex, race, color, handicap, religion, sexual orientation, or national and ethnic origin fell within its broader coverage. But this explicit notation, according to Grey, should have counted for, rather than against, the Policy because by making clear precisely what speech the general prohibition covered, the reference mitigated the potential chilling effect of the Policy on other expression.<sup>4</sup>

To evaluate this claim, it is necessary to take a step backward and ask what underlies the Court's distinction between direct and incidental restraints on expression.<sup>5</sup> The distinction makes no sense if what matters, under First Amendment doctrine, is the effects of a law on a speaker's expressive opportunities. The Stanford student who wishes to engage in race-based invective will "suffer" no more from a direct restriction on hate speech than from a generally applicable anti-discrimination regulation that covers all the speech affected by the direct restriction, but conduct in addition. The distinction likewise makes no sense if what matters is the effects of a law on an audience's ability to hear and consider a range of viewpoints. Again, the debate about race in the Stanford community will "suffer" no more from the one (speech-directed) form of regulation than from the other (generally applicable) kind. So much is always true of the distinction between direct and incidental restraints: the Court's use of the distinction cannot derive from considering the effects of such restraints, whether on a speaker or on an audience.<sup>6</sup>

---

<sup>4</sup> See Grey, *supra* note 3, at 923-24.

<sup>5</sup> For more expansive treatment of this subject, see Elena Kagan, *Private Speech, Public Purpose: The Role of Governmental Motive in First Amendment Analysis*, 63 U. CHI. L. REV. 413, 491-505 (1996); Frederick Schauer, *Cuban Cigars, Cuban Books, and the Problem of Incidental Restraints on Communications*, 26 WM. & MARY L. REV. 779 (1985); Geoffrey R. Stone, *Content-Neutral Restrictions*, 54 U. CHI. L. REV. 46, 105-14 (1987).

<sup>6</sup> To use a far-flung example, compare a (direct) law imposing a penny tax on the Sunday edition of the *New York Times* with a (generally applicable) law providing tax benefits for companies entering into certain kinds of mergers. Even if the effect of the direct law is nil and the effect of the generally applicable law is to restructure the whole communications industry, current doctrine subjects the former to strict scrutiny and the latter to mere rationality review.

But now assume that First Amendment law largely concerns motives, rather than effects — more specifically, that the doctrine has as its primary, though unstated, object the discovery of improper governmental motive.<sup>7</sup> This prohibited motive may roughly be termed “ideological”; it exists when simple disapproval of an idea — as distinct from a neutral evaluation of the harm that idea causes — enters into the decision to limit expression.<sup>8</sup> The Court, of course, cannot ascertain this illicit motive directly — or at least, cannot do so with any effectiveness. Hence, the Court (whether consciously or not is unimportant) has constructed and relied upon a set of rules and categories, most focusing on the facial aspects of a law, that operates as a proxy for this direct inquiry. These rules comprise tools to flush out impermissible motive and invalidate actions infected with it: they enforce the central command of the First Amendment that the government cannot interfere in the realm of speech simply because it finds some ideas correct and others abhorrent.

The doctrine of incidental restraints, as Grey himself recognizes,<sup>9</sup> serves precisely this function of assisting in the discovery of improper motive. A generally applicable law by definition targets not a particular idea, nor even ideas broadly speaking, but an object that need not, and usually does not, have any association with ideas whatsoever. The breadth of these laws makes them poor vehicles for censorial designs; they are instruments too blunt for either effecting or reflecting ideological disapproval of certain messages. (Consider, for example, the likelihood that a law prohibiting fires in public places — though encompassing such speech as the burning of an American flag — has resulted from ideological disapproval of certain messages.) Thus, incidental restrictions receive minimal constitutional scrutiny because of the likelihood that they will also be accidental restrictions in the relevant sense — that they will result from a process in which officials’ hostility toward ideas qua ideas played no role.

<sup>7</sup> For a broadscale defense of this proposition, discussing many aspects of First Amendment law, see Kagan, *supra* note 5.

<sup>8</sup> This definition of impermissible motive raises many hard questions, of both a conceptual and a practical nature. For discussion of these issues, which I cannot explore here, see generally *id.* at 428-37.

<sup>9</sup> See Grey, *supra* note 3, at 919.

With th  
cle and r  
not a req  
generally,  
where me  
tion that l  
nent part  
non-expre  
prohibitio  
crimination  
— all abo  
harms of  
mines the  
to opposit  
very oppo  
considered  
pression.

As a law  
cerns rise  
ously offer  
fires with  
ban to flag  
focused or  
speech —  
message. A  
idea — th  
decision-m  
Policy, mo  
exposition  
expression.  
den, the c  
doctrine’s  
tal restricti  
classify an

<sup>10</sup> See *id.* at

<sup>11</sup> Of cours  
be argued that  
obviously and i  
works well enou  
difficulty of find

With this as background, turn to the Appendix of Grey's article and review the text of the Stanford Policy.<sup>10</sup> The Policy is not a regulation that, in the manner of incidental restraints generally, refers to a broad class of activity, including but nowhere mentioning expression. The Policy is not even a regulation that breaks down a broad class of activity into all its component parts, listing expression but equivalently listing kinds of non-expressive conduct as falling within the scope of the general prohibition. The Policy, although referring to a broad anti-discrimination ideal, is nonetheless — on its face and by its terms — all about expression. It explicitly considers the benefits and harms of expression; weighs the one against the other; determines the point at which ideals of free inquiry should give way to opposing values. The Policy, in other words, constitutes the very opposite of the usual incidental restraint: a specific and considered judgment of the desirability of restricting certain expression.

As a law takes on this form, the Court's motive-based concerns rise to the fore. Consider, to continue the example previously offered, if a city were to replace its general ban on public fires with an ordinance explicitly discussing application of the ban to flag-burning. No one deciding whether to adopt the new, focused ordinance could do so without evaluating its effect on speech — more, without evaluating its effect on a particular message. And in considering this effect, sheer hostility of the idea — that is, impermissible motive — well might enter the decision-making process. So too when Stanford adopted its new Policy, moving from a generalized "morals code" to an explicit exposition of how this code applied to certain racist (sexist, etc.) expression. In general, as a limit on speech becomes less hidden, the danger of illicit motive increases: hence the current doctrine's distinction between facially direct and facially incidental restrictions.<sup>11</sup> For a court to do what Grey suggests — to classify an explicit speech-directed action as "incidental" whenever

---

<sup>10</sup> See *id.* at Appendix.

<sup>11</sup> Of course, this generalization, like all generalizations, sometimes fails; it even could be argued that it does not hold up in the Stanford case because the initial incidental ban obviously and importantly (even if not facially) applied to speech. But the generalization works well enough to make it a useful test for ascertaining governmental motive, given the difficulty of finding such motive directly.

er it can be conceptualized as a component of a broader, non-speech prohibition — would subvert the very basis of the doctrine. Such a move would prevent the doctrine of incidental restraints from performing its core function of ferreting out impermissible governmental motive.

Grey is right that the rule against directly referring to speech, if followed in this case, would have made the Policy's application to speech more vague and hence more chilling. But it is not surprising that First Amendment doctrine declines to take account of this point. First, the enhanced chilling effect that Grey notes is not usually, let alone invariably, the result of a narrow (i.e., the current) understanding of the category of incidental restraints. Such an effect arises here only because the contours of the general prohibition are unusually uncertain; in the more common case, a list of applications to speech will serve as much to confuse as to clarify the issue.<sup>12</sup> Second and more important, First Amendment doctrine, as I have suggested earlier, always cares less about effects than about motives.<sup>13</sup> In any clash between the two — in any case in which a concern with untoward effects points to one doctrinal rule and a concern with improper motive points to another — the doctrine tracks the concern with motive. The distinction between direct and incidental restraints, in both its broad outlines and its shadings, provides but a single instance.<sup>14</sup> Grey's attempt to rework the distinction — to divorce it from its underlying motive-based rationale, which in turn links it with the rest of First Amendment doctrine — thus was preordained for failure.

## II. CHALLENGING THE DOCTRINE OF INCIDENTAL RESTRAINTS

Perhaps recognizing the difficulty of labeling the Stanford policy an incidental rather than a direct restraint, Grey turns

<sup>12</sup> Consider, for example, the law against lighting fires in public places (incidentally restricting a person who burns a flag as a means of protest), or a law against vandalism (incidentally restricting a person who draws a swastika on a synagogue wall), or a law against trespass (incidentally restricting a person who burns a cross on private property). In cases of this kind — which are very much the norm — listing the law's potential applications to expression cannot serve a constitutionally legitimate purpose.

<sup>13</sup> See *supra* note 8 and accompanying text.

<sup>14</sup> See Kagan, *supra* note 5, at 491-505.

midway the distinction point is by cal race t goes, between speech co expression. I cant part l sage of gr ample, of a ban on t matic inju is nothing the Supre conflicts w ford Policy

In staki Antidiscrim can abstrac apparently fere, as civ sage attend communic cause iden phenomene exists all o well as do,

But it is the messag any other.

<sup>15</sup> See Grey,

<sup>16</sup> See, e.g.,

*Campus*, 1990 D

<sup>17</sup> Grey, *sup*

<sup>18</sup> 347 U.S.

<sup>19</sup> Converse

extreme version

UNMODIFIED: D

version, in part

CHI. L. REV. 79!

midway through his article to challenging the coherence of that distinction, at least when civil rights law is at issue.<sup>15</sup> The basic point is by now familiar, having become a staple of certain critical race theory.<sup>16</sup> We cannot distinguish, or so the argument goes, between civil rights statutes (incidental restraints) and hate speech codes (direct restraints), because both really target expression. In Grey's words, "we prohibit discrimination in significant part because of its 'expressive content,' because of the message of group inferiority it sends."<sup>17</sup> The proscription, for example, of segregated schools should be viewed at least in part as a ban on the message of racial inferiority, deemed to cause stigmatic injury. The proscription contained in a hate speech code is nothing more. Hence, to put the point in its bluntest form, the Supreme Court's decision in *Brown v. Board of Education*<sup>18</sup> conflicts with the district court's decision invalidating the Stanford Policy.

In staking this claim, Grey no doubt is on to something. Antidiscrimination laws are in part about message. Indeed, we can abstract Grey's point, because so too are other kinds of laws apparently directed at conduct. Many incidental restraints interfere, as civil rights laws do, with the communication of a message attending an act, as well as the injury that follows from that communication. This is because both conduct and speech may cause identical "expressive" harms, such as stigmatization. The phenomenon is not limited to the sphere of civil rights, but exists all over, by virtue of the simple fact that most acts say, as well as do, something.<sup>19</sup>

But it is well not to overstate the equivalence of an act and the message it carries, whether in the field of civil rights or in any other. Grey provides, though perhaps does not highlight

---

<sup>15</sup> See Grey, *supra* note 3, at 934.

<sup>16</sup> See, e.g., Charles R. Lawrence III, *If He Hollers Let Him Go: Regulating Racist Speech on Campus*, 1990 DUKE L.J. 431, 449-57.

<sup>17</sup> Grey, *supra* note 3, at 934.

<sup>18</sup> 347 U.S. 483 (1954).

<sup>19</sup> Conversely, most speech does as well as says something in some sense. For the most extreme version of this claim and its implications, see CATHARINE A. MACKINNON, *FEMINISM UNMODIFIED: DISCOURSES ON LIFE AND LAW* 129-30, 193-94 (1987). For a more moderate version, in part critiquing MacKinnon, see Cass R. Sunstein, *Words, Conduct, Caste*, 60 U. CHI. L. REV. 795, 836-40 (1993).

sufficiently, the appropriate caveat: after all, he notes, discrimination (in employment, housing, or other material benefit) remains discrimination even when well hidden.<sup>20</sup> Message matters, but it is not all that matters; when the government forbids, say, segregated schools, it does more than shape the world of communication. This wider significance is precisely what justifies the generalization, discussed earlier, that an incidental restriction is less likely than a direct restriction to arise from hostility toward certain messages: because the government is regulating on the basis of something other, or at least more, than expressive content, this illicit factor should have less effect on the decision-making process.

Perhaps more important, I count Grey's claim as a prime example of a category of academic ideas that I call Ultimately Useless Insights — ideas that, however true and even important in some sense, do not and cannot assist in the elaboration of legal doctrine. Grey himself half-concedes this point by noting the logical conclusion of his insight: If civil rights laws partly target the "stigmatic messages" associated with conduct and if, therefore, the same messages, when conveyed by speech, are likewise subject to limit, "there wouldn't," in Grey's own words, "be much to freedom of speech on some of the central contested issues in our politics and culture."<sup>21</sup> Under the proposed analysis, the government (or a university operating under the government's rules) could restrict not only race-based (or sex-based, etc.) fighting words, but all speech that stigmatizes on the basis of group characteristics. The care that Grey put into crafting a carefully limited restriction, applying only to fighting words, would have been wasted. The expressive content of the conduct that civil rights laws target would render vast amounts of speech on race (or gender, etc.) proscribable.

The same point applies generally. If the conduct encompassed by an incidental restriction has some expressive content, as almost all conduct does, Grey's insight would seem to allow direct

<sup>20</sup> See Grey, *supra* note 3, at 934-36.

<sup>21</sup> *Id.* at 937. The alternative conclusion of Grey's insight is that there wouldn't be much to civil rights laws. This conclusion would hold if the message associated with discriminatory conduct brought laws prohibiting that conduct under the protection of the First Amendment.

restriction though G require th that is, of sis becom speech, o distinction whether t law, thus :

Grey m conceptua much to only certa critical ins minds" arg judicially : then, *Brou*

### III.

Stanford — a polic the Funda ty, person: case-by-cas prohibit, i "harassme harassmen regulation,

<sup>22</sup> *Id.* at 89

<sup>23</sup> *Id.* at 89

<sup>24</sup> To say t ment is irrelev. scrutiny (usual 391 U.S. 367. : have to meet c dictum in R.A. case where the threats or fight cation of an ir

restriction of any speech with the same message. Alternatively, though Grey does not consider the possibility, his insight might require the protection of any conduct expressing a message — that is, of conduct generally. Either way, First Amendment analysis becomes impossible: either the First Amendment protects no speech, or it protects speech and all else in addition. Some distinction between direct and incidental restraints, regardless whether the precise motive-related distinction used in current law, thus seems a necessary component of a free speech system.

Grey may agree with this much; perhaps in questioning the conceptual foundations of the distinction, he wishes not so much to overturn it as to render it irrelevant to certain (but only certain) civil rights-type cases. But if that is the point of his critical insight, he must show how what he calls the “hearts and minds” argument can fit within, rather than subvert, a workable, judicially administrable doctrine of incidental restrictions. Until then, *Brown* will not justify the Stanford Policy.

### III. POLITICS, THE POLICY, AND THE DOCTRINE OF INCIDENTAL RESTRAINTS

Stanford, of course, had a policy before (and after) the Policy — a policy that the Policy was supposed to enhance. Termed the Fundamental Standard, it requires “respect for order, morality, personal honor and the rights of others.”<sup>22</sup> Interpreted on a case-by-case basis over the years, the Standard is understood to prohibit, in the words of the President of the University, all “harassment, whether accompanied by speech or not, including harassment that is motivated by racial or other bigotry.”<sup>23</sup> This regulation, unlike Grey’s Policy, is an incidental restraint.<sup>24</sup>

---

<sup>22</sup> *Id.* at 893 n.6 (quoting Stanford’s Fundamental Standard).

<sup>23</sup> *Id.* at 897 n.20 (quoting Stanford President Gerhard Casper).

<sup>24</sup> To say that the Standard is an incidental restraint is not to say that the First Amendment is irrelevant. An incidental restraint, when applied to speech, may trigger heightened scrutiny (usually of an intermediate level), as the seminal case of *United States v. O’Brien*, 391 U.S. 367, 376-77 (1968), shows. Applications of the Standard to expression thus may have to meet certain First Amendment requirements. But I agree with Grey — and with the dictum in *R.A.V. v. City of St. Paul*, 505 U.S. 377, 389 (1992) — that this would not be the case where the speech affected falls within a category of wholly proscribable speech, as do threats or fighting words. And even when speech is fully protected, as in *O’Brien*, the application of an incidental restriction to the speech usually (though not always) will receive

Like many incidental restraints, the Standard has a potentially profound effect on expression. The Standard, as interpreted, already may have prohibited all of the speech specifically barred by the Policy. No doubt the Standard prohibited more speech besides. Judged solely by its efficacy in eradicating a certain kind of harmful speech, the direct restriction held no advantage over the incidental restraint.

Proponents of the Policy might claim for it a symbolic function. True, the Standard might succeed in punishing bigoted speech of a harassing nature. What the Standard cannot do — precisely because it is an incidental restriction — is to send a clear message about the University's attitude toward this expression. Grey has argued in support of his Policy on another occasion that it was necessary to convey the University's attitude toward bigotry and intolerance.<sup>25</sup> Similarly, Richard Delgado has urged on behalf of his proposed tort action for racial insults, which Grey approves, that it "communicat[es] to the perpetrator and to society that such abuse will not be tolerated."<sup>26</sup> The general proscription can accomplish all the garden-variety ends of regulation; the particular, speech-directed proscription is needed, or so the argument runs, to communicate as forcefully as possible the governmental actor's commitment to the goal of equality.

This understanding of the Policy, which views an orientation toward speech as critical to the achievement of the regulatory goal, itself casts doubt on Grey's claim to have drafted an incidental restriction. Indeed, this view of the Policy, by highlighting the different motives that may lie behind direct and incidental restrictions, suggests one of the key reasons for distinguishing between these kinds of regulation. But I want to end this commentary by placing these doctrinal issues to one side and evaluating Grey's handiwork solely in terms of its own primary objective: the advancement of equality in the University and the broader community. This evaluation suggests some practical

---

more deferential treatment than a direct restraint on the same expression.

<sup>25</sup> See Thomas C. Grey, *Civil Rights vs. Civil Liberties: The Case of Discriminatory Verbal Harassment*, 8 SOC. PHIL. & POL'Y 81, 104 (Spring 1991) (writing that "I concede that the main purposes behind the proposal are in a certain sense educative or symbolic.").

<sup>26</sup> Richard Delgado, *Words That Wound: A Tort Action for Racial Insults, Epithets, and Name Calling*, 17 HARV. C.R.-C.L. L. REV. 133, 147 (1982).

political drawbacks of moving, as Grey and Stanford decided to do, from the generally applicable to the speech directed.

Grey himself alludes to such concerns, in the conclusion to his article, when he discusses the way in which adoption of the Stanford Policy distracted from debate, and potential progress, on more important issues of race and gender.<sup>27</sup> Grey notes that a broader argument about affirmative action on the Stanford campus was diverted into the controversy over fighting words. And citing Henry Louis Gates's potent arguments, Grey more generally concedes the ability of disputes on speech to shift attention from, even excuse inattention to, weightier issues, extending far beyond the academic setting, of inequality in housing, employment, and other material goods.<sup>28</sup> But even while acknowledging these costs, Grey stubbornly hangs on to the Stanford Policy, just as other academics in other educational institutions insist on still broader restrictions on expression. Hence occurs the direction of energy away from the alleviation of material inequalities and toward the elimination — yes, of “only words”<sup>29</sup> — of “insults, epithets, and name calling.”<sup>30</sup>

The costs of opening this two-front war are higher even than in the usual case — greater than the inevitable loss of focus and dispersion of resources. As an initial matter, the second front here occurs in the one place where the opposition — however disingenuous and hypocritical in fact — seems to many to hold the high ground.<sup>31</sup> It is poor strategy to turn a battle about discrimination into a battle about speech — to mount the kind of attack most likely to transform the forces of hatred into the

---

<sup>27</sup> Grey, *supra* note 3, at 939-45.

<sup>28</sup> See *id.* at 928. Gates terms the critical race theorists' focus on hate speech “a see-no-evil, hear-no-evil approach toward racial inequality,” noting that “even if hate [speech] did disappear, aggregative patterns of segregation and segmentation in housing and employment would not disappear.” Henry L. Gates, Jr., *Let Them Talk: Why Civil Liberties Pose No Threat to Civil Rights*, THE NEW REPUBLIC, Sept. 20, 1993, at 49.

<sup>29</sup> CATHARINE A. MACKINNON, ONLY WORDS (1987).

<sup>30</sup> See generally Delgado, *supra* note 26.

<sup>31</sup> Even Charles Lawrence, a defender of at least some speech codes, has noted:

I fear that by framing the debate as we have — as one in which the liberty of free speech is in conflict with the elimination of racism — we have advanced the cause of racial oppression and . . . placed the bigot on the moral high ground, fanning the rising flames of racism.

Lawrence, *supra* note 16, at 436.

defenders of constitutional liberty. Relatedly, the second front here causes not merely the division, but the permanent loss of resources. As speech codes, in Grey's words, "set civil rights advocates and civil libertarians . . . against each other," they threaten to rend the coalitions that have served well on other, more important issues.<sup>32</sup> Grey's tactic of limiting and hedging such a code can contain, but not avert, this damage.

I suspect that the temptation to fight on this ground, seemingly irrespective of tactical advantage, derives from frustration, even desperation, over the slow pace of progress in eradicating the tangible, socio-economic inequalities existing between blacks and whites and, to a lesser extent, between men and women. The magnitude and duration of these inequalities may make them appear impervious to political (let alone to academic) efforts. We do not know how to solve these problems; we may not even know how (or perhaps we are afraid) to talk about them. So some succumb to the allure of sideshows, such as the one involving the Stanford Policy. There, the issues seem contained, the solutions discernible, the link between activism and result still full of potential. Victory is achievable, if ultimately empty.<sup>33</sup>

The lesson the Stanford experience suggests to me is one about resisting such urges. If, as Grey laments, "the effort ended up with a grotesquely unreal portrayal of Stanford as a campus under the dominion of the thought police"<sup>34</sup> — if in doing so, the effort only undermined serious attempts to advance the goal of equality — neither Grey nor Stanford should profess much surprise. Stanford's course of action — its shift from a generally applicable ban on harassment, including racial or sexual harassment, whether or not accompanied by expression, to a targeted ban on certain bigoted harassing speech — misjudged the political, as well as the legal, environment. Just as the Policy, in directly rather than incidentally restricting speech, became vulnerable to judicial invalidation, so too did it become a focal point

---

<sup>32</sup> Grey, *supra* note 3, at 944-45.

<sup>33</sup> See Gates, *supra* note 28, at 49 (stating that "[t]he advocates of speech restrictions will grow disenchanted not with their failures, but with their victories, and the movement will come to seem yet another curious byway in the long history of our racial desperation").

<sup>34</sup> Grey, *supra* note 3, at 939-40.

for all manner of public complaint over Stanford's race and gender policies. The law and the politics of moving from the general to the particular thus coincided. From either perspective, Stanford and Professor Grey should have declined to convert an incidental into a direct restraint.

# The University of Chicago Law Review



---

Confirmation Messes, Old and New

*Elena Kagan*

---

Volume 62

Spring 1995

Number 2

## REVIEW

### Confirmation Messes, Old and New

*Elena Kagan*†

*The Confirmation Mess.* Stephen L. Carter.  
Basic Books, 1994. Pp xiii, 252.

What confirmation mess?

Stephen Carter's new book decries the state of the confirmation process, especially for Supreme Court nominees. "The confirmation mess," in Carter's (noninterrogatory) phrase, consists of both the brutalization and the politicization of the process by which the nation selects its highest judges. That process, Carter insists, is replete with meanness, dishonesty, and distortion. More, and worse, it demands of nominees that they reveal their views on important legal issues, thus threatening to limit the Court "to people who have adequately demonstrated their closed-mindedness" (p xi). A misguided focus on the results of controversial cases and on the probable voting patterns of would-be Justices, Carter argues, produces a noxious and destructive process. Carter's paradigm case, almost needless to say, is the failed nomination of Robert Bork.

But to observers of more recent nominations to the Supreme Court, Carter's description must seem antiquated. President

---

† Assistant Professor of Law, University of Chicago. I served as Special Counsel to the United States Senate Committee on the Judiciary in connection with the nomination of Justice Ruth Bader Ginsburg to the Supreme Court. The views expressed in this Review are, of course, mine alone. I thank Al Alschuler, David Currie, Richard Epstein, Tracy Meares, Cass Sunstein, and Mark Tushnet for helpful comments. The Class of 1964 Fund and the Russell J. Parsons Faculty Research Fund provided financial support.

Clinton's nominees, then-Judges Ruth Bader Ginsburg and Stephen Breyer, confronted no unfair or nasty opposition; to the contrary, their confirmation hearings became official lovefests. More important, both nominees felt free to decline to disclose their views on controversial issues and cases. They stonewalled the Judiciary Committee to great effect, as senators greeted their "nonanswer" answers with equanimity and resigned good humor. And even before the confirmation process became quite so cozy (which is to say, even before the turn toward nominating well-known and well-respected moderates), the practice to which Carter most objects—the discussion of a nominee's views on legal issues—had almost completely lapsed. Justices Kennedy, Souter, and Thomas, no less than Justices Ginsburg and Breyer, rebuffed all attempts to explore their opinions of important principles and cases. Professor Carter, it seems, wrote his book too late. Where, today, is the confirmation mess he laments?

The recent hearings on Supreme Court nominees, though, suggest another question: might we now have a distinct and more troubling confirmation mess? If recent hearings lacked acrimony, they also lacked seriousness and substance. The problem was the opposite of what Carter describes: not that the Senate focused too much on a nominee's legal views, but that it did so far too little. Otherwise put, the current "confirmation mess" derives not from the role the Senate assumed in evaluating Judge Bork, but from the Senate's subsequent abandonment of that role and function. When the Senate ceases to engage nominees in meaningful discussion of legal issues, the confirmation process takes on an air of vacuity and farce, and the Senate becomes incapable of either properly evaluating nominees or appropriately educating the public. Whatever imperfections may have attended the Bork hearings pale in comparison with these recent failures. Out, then, with the new mess and in with the old!<sup>1</sup>

#### I. CARTER'S CRITIQUE

Carter depicts a confirmation process out of control—a process in which we attend to the wrong things in the wrong manner, in which we abjure reasoned dialogue about qualifications in favor of hysterical rantings about personalities and politics. Car-

---

<sup>1</sup> And no, I haven't changed my mind since, several months after I drafted this Review, the Senate turned Republican and Orrin Hatch assumed the chairmanship of the Judiciary Committee. The conclusion of this Review still holds—even if I am no longer quite so sanguine about it.

ter is no partisan in this description; he blames Republicans and Democrats, right and left alike (pp 10, 142). Similarly, Carter takes no sides as between the President and the Senate; he assumes that both ought to evaluate judicial candidates by the same criteria and argues that both have performed poorly this evaluative function (pp 29-30). Carter views the current mess as having deep roots. He refers often to the attempt of segregationist senators to defeat the nomination of Thurgood Marshall (pp 62-63) and describes as well some yet more distant confirmation battles (pp 65-73). Although he focuses on the nomination and confirmation of Supreme Court Justices, he buttresses his case with discussion of the recent travails of Lani Guinier (pp 37-44) and Zoe Baird (pp 25-28). Always, though, the face in the foreground is Robert Bork's. Carter's understanding of the Bork hearings informs—sometimes explicitly, sometimes not—the whole of his argument and analysis.

Carter identifies two cardinal flaws in the confirmation process. The first concerns the absence of "honesty" and "decency" (p ix). Here Carter laments the deterioration of public debate over nominations into "the intellectual equivalent of a barroom brawl" (p x). He catalogues the ways in which opponents demonize nominees and distort their records, referring to the many apparently purposeful misreadings of the writings of Robert Bork (pp 45-52) and Lani Guinier (pp 39-44). He describes the avid search for disqualifying factors, whether of a personal kind (for example, illegal nannies) or of a professional nature (for example, ill-conceived footnotes in scholarly articles) (pp 25, 42-43). He deplors "smears" and "soundbites" (p 206)—the way in which media coverage turns nominations into extravaganzas, the extent to which public relations strategy becomes all-important. And in a semi-mystical manner, he castigates our refusal to forgive sin, accept redemption, and acknowledge the complexity of human beings, including those nominated to high office (pp 183-84).

The second vice of the confirmation process, according to Carter, lies in its focus on a nominee's probable future voting record. In Carter's portrayal, the President, Senate, press, interest groups, and public all evaluate nominees primarily by plumbing their views on controversial legal issues, such as the death penalty or abortion (pp 54-56). Carter's paradigmatic case, again, is Robert Bork, a judge of superior objective qualifications whose views on constitutional method and issues led to the defeat of his nomination. Carter is "struck" by the failure of participants in the Bork hearings to consider "that trying to get him to tell the

nation how he would vote on controversial cases if confirmed might pose a greater long-run danger to the Republic than confirming him" (p x). This danger, Carter avers, arises from the damage such inquiry does to judicial independence. Examination of a nominee's views on contested constitutional matters, Carter claims, gives the public too great a chance to influence how the judiciary will decide these issues, precisely by enabling the public to reject a nominee on grounds of substance (p 115). At the same time, such inquiry undermines the eventual Justice's ability (and the public's belief in the Justice's ability) to decide cases impartially, based on the facts at issue and the arguments presented, rather than on the Justice's prior views or commitments (p 56).

The failures of the confirmation process, Carter urges, ultimately have less to do with rules and procedures than with public "attitudes"—specifically, "our attitudes toward the Court as an institution and the work it does for the society" (p 188). We view the Court as a dispenser of decisions—as to individual cases of course, but also as to hotly disputed public issues. Our evaluation of the Court coincides with our evaluation of the results it reaches (p 57). Because we see the Court in terms of results, we yearn to pack it with Justices who will always arrive at the "right" decisions. And because the decisions of the Court indeed have consequence, we feel justified, as we pursue this project, in resorting to "shameless exaggeration" and misleading rhetoric (p 51). The key to change, according to Carter, lies in viewing the Court in a different—a more "mundane and lawyerly"—manner (p 206). And although Carter is unclear on the point, this seems to mean judging the Court less in terms of the results it reaches than in terms of its level of skill and craftsmanship.

In keeping with this analysis, Carter advocates a return to confirmation proceedings that focus on a nominee's technical qualifications—in other words, his legal aptitude, skills, and experience (pp 161-62). At times, Carter suggests that this set of qualifications constitutes the only proper criterion of judgment (pp 187-88). But Carter in the end draws back from this position, which he admits would provide no lever to oppose a nominee, otherwise qualified, who wished to overturn a case like *Brown v Board*<sup>2</sup> (pp 119-21). Carter urges, as a safeguard against extremism of this kind, an inquiry into whether a nominee subscribes to the "firm moral consensus" of society (p 121). The Senate, Carter writes, should resolve this question by "undertak[ing] moral

---

<sup>2</sup> *Brown v Board of Education*, 347 US 483 (1954).

inquiry, both into the world view of the nominee and, if necessary, into the nominee's conduct" (p 124). This inquiry, in other words, would involve a determination of whether a nominee has the "right moral instincts" and whether his "personal moral decisions seem generally sound" (p 152). Carter views this inquiry as wholly distinct from an approach that asks about a nominee's legal views or philosophy (id). He suggests, for example, that the Senate ask a nominee not whether discriminatory private clubs violate the Constitution, but whether "the nominee has belonged to a club with such policies" (id). An assessment of moral judgment alone, independent of legal judgment, would combine with an evaluation of legal aptitude to form Carter's ideal confirmation process.

## II. CURRENT EVENTS

Does Carter's critique of the confirmation process ring true? It might have done so eight years ago. It ought not to do so now.

Carter tries to update his book, to make it more than a comment on the Bork proceedings. He invokes the nomination, eventually withdrawn, of Lani Guinier to serve as Assistant Attorney General for Civil Rights (pp 37-44). Consider, Carter implores us, the distortion of Guinier's academic work, initially by her many enemies, finally and fatally by some she thought friends. Do not the exaggeration, name-calling, and hyperbole that surrounded the discussion of Guinier's views prove the existence of a confirmation mess? And Carter then invokes the battle over the nomination of Clarence Thomas to serve as a Supreme Court Justice (pp 138-42). Recall, Carter tells us (and it is not hard to do), the intensity and wrath surrounding that battle—the fury with which the partisans of Thomas and Anita Hill, respectively, exchanged charge and countercharge and bloodied previously unsullied reputations. Does not this episode, this display of raw emotion and this unrelenting focus on personal traits and behavior, demonstrate again the existence of a confirmation mess?

Well, no—not on either count, at least if the term "confirmation mess" signifies a problem both specific to and common among confirmation battles. Carter is right to note the distortions in the debate over Guinier's prior writings; but he is wrong to think they derived from a special attribute of the confirmation process. It is unfortunate but true that distortions of this kind mar public debate on all important issues. Professor Carter, meet Harry and Louise; they may convince you that the Guinier episode is less a part of a confirmation mess than of a government

mess, the sources and effects of which lie well beyond your book's purview. And the Thomas incident, proposed as exemplar or parable, suffers from the converse flaw. That incident is unique among confirmation hearings and, with any reasonable amount of luck, will remain so. The way the Senate handled confidential charges of a devastating nature on a subject at a fault line of contemporary culture reveals very little about the broader confirmation process.

Indeed, Carter's essential critique of the confirmation process—that it focuses too much on the nominee's views on disputed legal issues—applies neither to the Guinier episode nor to the Thomas hearings. Carter concedes that the Senate ought to inquire into the views and policies of nominees to the executive branch, for whom "independence" is no virtue (p 32). The public debate over Guinier's articles (problems of distortion to one side) thus fails to implicate Carter's concern with the focus of the process on legal issues. And so too of the Thomas hearings. Carter's own description of the "mess" surrounding that nomination highlights the Senate's inquiry into the charges of sexual harassment and not its investigation of the nominee's legal opinions (pp 133-45). The emphasis is not surprising. No one can remember the portion of the hearings devoted to Justice Thomas's legal views, and for good reason: Justice Thomas, or so he assured us, already had "stripped down like a runner" and so had none to speak of.<sup>3</sup> The apparent "mess" of the Thomas hearings thus arose not from the exploration of legal philosophy that Carter abjures, but instead from the inquiry into moral practice and principle that he recommends to the Senate as an alternative.<sup>4</sup>

What, then, of the "confirmation mess" as Carter defines it—the threat to judicial independence resulting from a misplaced focus on the nominee's legal views and philosophy? Lacking support for his argument in the recent controversies surrounding Guinier and Thomas, Carter must recede to the Bork hearings for a paradigm. But time has overtaken this illustration: no subsequent nomination fits Carter's Bork-based model

---

<sup>3</sup> Clarence Thomas, as quoted in Linda Greenhouse, *The Thomas Hearings: In Trying to Clarify What He Is Not, Thomas Opens Questions of What He Is*, NY Times A19 (Sept 13, 1991).

<sup>4</sup> The same is true of the controversy surrounding the nomination of Zoe Baird as Attorney General. As Carter discusses, Baird's nomination ran into trouble because she had hired illegal immigrants and then failed to pay social security taxes on their salaries (pp 25-28). Here, too, the dispute arose from an inquiry into the nominee's personal conduct, rather than her views and policies.

any better than do the nominations of Guinier or Thomas. Not since Bork (as Carter himself admits) has any nominee candidly discussed, or felt a need to discuss, his or her views and philosophy (pp 57-59). It is true that in recent hearings senators of all stripes have proclaimed their prerogative to explore a nominee's approach to constitutional problems. The idea of substantive inquiry is accepted today to a far greater extent than it was a decade ago.<sup>5</sup> But the practice of substantive inquiry has suffered a precipitous fall since the Bork hearings, so much so that today it hardly deserves the title "practice" at all. To demonstrate this point, it is only necessary to review the recent hearings of Ruth Bader Ginsburg and Stephen Breyer—one occurring before, the other after, publication of Carter's book. Consider the way these then-judges addressed issues of substance and then ask of what Carter's "confirmation mess" in truth consists.

Justice Ginsburg's favored technique took the form of a pincer movement. When asked a specific question on a constitutional issue, Ginsburg replied (along Carter's favored lines) that an answer might forecast a vote and thus contravene the norm of judicial impartiality. Said Ginsburg: "I think when you ask me about specific cases, I have to say that I am not going to give an advisory opinion on any specific scenario, because . . . that scenario might come before me."<sup>6</sup> But when asked a more general question, Ginsburg replied that a judge could deal in specifics only; abstractions, even hypotheticals, took the good judge beyond her calling. Again said Ginsburg: "I prefer not to . . . talk in grand terms about principles that have to be applied in concrete cases. I like to reason from the specific case."<sup>7</sup> Some room may have remained in theory between these two responses; perhaps a senator could learn something about Justice Ginsburg's legal

---

<sup>5</sup> Senator Joseph Biden made this point near the beginning of the Ginsburg hearings. After listening, in turn, to Senators Hatch, Kennedy, Metzenbaum, and Simpson expound on the need to question the nominee about her judicial philosophy, Senator Biden said: "I might note it is remarkable that seven years ago the hearing we had here was somewhat more controversial, and I made a speech that mentioned the 'p' word, philosophy, that we should examine the philosophy, and most . . . said that was not appropriate. At least we have crossed that hurdle. No one is arguing that anymore." *Nomination of Ruth Bader Ginsburg, to be Associate Justice of the Supreme Court of the United States, Hearings before the Senate Committee on the Judiciary, 103d Cong, 1st Sess 21 (July 20-23, 1993) ("Confirmation Hearings for Ginsburg")*.

<sup>6</sup> *Id.* at 184.

<sup>7</sup> *Id.* at 180. See also *id.* at 333 ("I can't answer an abstract issue. I work from a specific case based on the record of that case, the briefs that are presented, the parties' presentations, and decide the case in light of that record, those briefs. I simply cannot, even in areas that I know very well, answer an issue abstracted from a concrete case.").

views if he pitched his question at precisely the right level of generality. But in practice, the potential gap closed to a sliver given Ginsburg's understanding of what counted as "too specific" (roughly, anything that might have some bearing on a case that might some day come before the Court) and what counted as "too general" (roughly, anything else worthy of mention).

So, for example, in a colloquy with Senator Feinstein on the Second Amendment, Ginsburg first confronted the question whether she agreed with a fifty-four-year-old Supreme Court precedent<sup>8</sup> on the subject and with the interpretation that lower courts unanimously had given it. Replied Ginsburg: "The last time the Supreme Court spoke to this question was 1939. You summarized what that was, and you also summarized the state of law in the lower courts. But this is a question that may well be before the Court again . . . and because of where I sit it would be inappropriate for me to say anything more than that."<sup>9</sup> The Senator continued: if the Judge could not discuss a particular case, even one decided fifty years ago, could the Judge say something about "the methodology [she] might apply" and "the factors [she] might look at" in determining the validity of that case or the meaning of the Second Amendment?<sup>10</sup> "I wish I could Senator," Ginsburg replied, "but . . . apart from the specific context I really can't expound on it."<sup>11</sup> "Why not?" the Senator might have asked. Because the question functioned at too high a level of abstraction: "I would have to consider, as I have said many times today, the specific case, the briefs and the arguments that would be made."<sup>12</sup> Many times indeed. So concluded a typical exchange in the confirmation hearing of Justice Ginsburg.

Justice Breyer was smoother than Justice Ginsburg, but ultimately no more forthcoming. His favored approach was the "grey area" test: if a question fell within this area—if it asked him to comment on issues not yet definitively closed (and therefore still a matter of interest)—he must, he said, decline to comment.<sup>13</sup> Like Justice Ginsburg, he could provide personal anecd-

---

<sup>8</sup> *United States v Miller*, 307 US 174 (1939).

<sup>9</sup> Confirmation Hearings for Ginsburg at 241-42 (cited in note 5).

<sup>10</sup> *Id.* at 242.

<sup>11</sup> *Id.*

<sup>12</sup> *Id.*

<sup>13</sup> Confirmation Hearings for Stephen G. Breyer to be an Associate Justice of the United States Supreme Court, Senate Committee on the Judiciary, 103d Cong, 2d Sess 85 (July 12, 1994) (Miller Reporting transcript). Sometimes Justice Breyer referred to this test as the "up in the air" test. So, for example, when Chairman Biden asked him to comment on the burden imposed on the government to sustain economic regulation,

dotes—the relevance of which were open to question. He could state settled law—but not whether he agreed with the settlement. He could explain the importance and difficulty of a legal issue—without suggesting which important and difficult resolution he favored. What he could not do was to respond directly to questions regarding his legal positions. Throughout his testimony, Breyer refused to answer not merely questions concerning pending cases, but questions relating in any way to any issue that the Supreme Court might one day face.

I do not mean to overstate the case; Justice Ginsburg and Justice Breyer did provide snippets of information. Both Justices discussed with candor and enthusiasm issues on which they previously had written. So the Judiciary Committee and public alike learned much about Justice Ginsburg's current views on gender discrimination and abortion and about Justice Breyer's thoughts on regulatory policy. Both Justices, too, allowed an occasional glimpse of what might be termed, with some slight exaggeration, a judicial philosophy. A close observer of the hearings thus might have made a quick sketch of Justice Ginsburg as a cautious, incrementalist common lawyer and of Justice Breyer as an antiformalist problem solver. (But how much of this sketch in fact would have derived from preconceptions of the Justices, based on their judicial opinions and scholarly articles?) If most of the testimony disclosed only the insignificant and the obvious—did anyone need to hear on no less than three separate occasions that Justice Ginsburg disagreed with *Dred Scott*?<sup>14</sup>—a small portion revealed something of the nominee's conception of judging.

Neither do I mean to deride Justices Ginsburg and Breyer for the approach each took to testifying. I am sure each believed (along with Carter) that disclosing his or her views on legal issues threatened the independence of the judiciary. (It is a view, I suspect, which for obvious reasons is highly correlated with membership in the third branch of government.<sup>15</sup>) More, I am sure

---

Breyer noted that "this is a matter . . . still up in the air." When the Chairman replied "[t]hat is why I am trying to get you to talk about it, because you may bring it down to the ground," Justice Breyer repeated that "I have a problem talking about things that are up in the air." *Id.* at 55 (July 12, 1994).

<sup>14</sup> *Dred Scott v. Sandford*, 60 US 393 (1856). See, for example, Confirmation Hearings for Ginsburg at 126, 188, 270 (cited in note 5).

<sup>15</sup> In 1959, lawyer William Rehnquist wrote an article criticizing the Senate's consideration of the nomination of Charles Evans Whittaker to the Supreme Court. The Senate, he stated, had "succeeded in adducing only the following facts: . . . proceeds from skunk trapping in rural Kansas assisted him in obtaining his early education; . . . he was the

both judges knew that they were playing the game in full accordance with a set of rules that others had established before them. If most prior nominees have avoided disclosing their views on legal issues, it is hard to fault Justice Ginsburg or Justice Breyer for declining to proffer this information. And finally, I suspect that both appreciated that, for them (as for most), the safest and surest route to the prize lay in alternating platitudinous statement and judicious silence. Who would have done anything different, in the absence of pressure from members of Congress?

And of such pressure, there was little evidence. To be sure, an occasional senator complained of the dearth of substantive comment, most vocally during the preternaturally controlled testimony of Justice Ginsburg. Chairman Biden and Senator Spector in particular expressed impatience with the game as played. Spector warned that the Judiciary Committee one day would "rear up on its hind legs" and reject a nominee who refused to answer questions, for that reason only (p 54). And Biden lamented that no "nominee would ever satisfy me in terms of being as expansive about their views as I would like."<sup>16</sup> But for the most part, the senators acceded to the reticence of the nominees before them with good grace and humor. Senator Simon sympathetically commented to Justice Breyer: "You are in a situation today . . . where you do not want to offend any of us, and I understand that. I hope the time will come when you may think it appropriate . . . to speak out on this issue."<sup>17</sup> Senator DeConcini similarly remarked to Justice Ginsburg that it was "fun" and "intellectually challenging"—a sort of chess game in real life—for a senator to "try[ ] to get inside the mind of a nominee . . . without violating their oath and their potential conflicts . . . ."<sup>18</sup> And of course no one voted against either nominee

---

first Missourian ever appointed to the Supreme Court; [and] since he had been born in Kansas but now resided in Missouri, his nomination honored two states." William Rehnquist, *The Making of a Supreme Court Justice*, Harv L Rec 7, 8 (Oct 8, 1959). Rehnquist specifically complained about the Senate's failure to ask Justice Whittaker about his views on equal protection and due process. *Id.* at 10. By 1986, when he appeared before the Senate Judiciary Committee as a sitting Associate Justice and a nominee for Chief Justice, Rehnquist had changed his mind about the propriety of such inquiries.

<sup>16</sup> Confirmation Hearings for Ginsburg at 259 (cited in note 5). In a similar vein, Senator Cohen accused Justice Ginsburg of resorting to "delphic ambiguity" in her responses. Senator Cohen recalled the story of the general who asked the oracle what would occur if he (the general) invaded Greece. When the oracle responded that a great army would fall, the general mounted the invasion—only to discover that the great army to which the oracle had referred was his own. See *id.* at 220.

<sup>17</sup> Confirmation Hearings for Breyer at 77-78 (July 13, 1994) (cited in note 13).

<sup>18</sup> Confirmation Hearings for Ginsburg at 330 (cited in note 5).

CLINTON LIBRARY PHOTOCOPY

on the ground that he or she had declined to answer questions relating to important legal issues.

The ease of these proceedings in part reflected the nature of both the nominations and the political context. First replace divided government with single-party control of the White House and Senate. Now posit a President with an ambitious legislative agenda, requiring him to retain support in Congress, but with no judicial agenda to speak of.<sup>19</sup> Assume, as a result, that this President nominates two clear moderates, known and trusted by leading senators of both the majority and the minority parties. Throw in that each nominee is a person of extraordinary ability and distinction. Finally, add that the Court's rulings on some of the hot-button issues of recent times—most notably abortion, but also school prayer and the death penalty—today seem relatively stable. This is a recipe—now proved successful—for confirmation order, exactly opposite to the state of anarchy depicted by Carter. At the least, this suggests what David Strauss has argued in another review of Carter's book:<sup>20</sup> that the culprit in Carter's story is nothing so grand and seemingly timeless as the American public's attitudes toward the courts; that the cause of Carter's "mess" is the simple attempt of the Reagan and Bush administrations to impose an ideologically charged vision of the judiciary in an unsympathetic political climate.

But even this view overstates the longevity of the "confirmation mess," as Carter defines it. That so-called mess in fact ended long before President Clinton's nominations; it ended right after it began, with the defeat of the nomination of Robert Bork. The Senate overwhelmingly approved the nominations of Justices Kennedy and Souter after they gave testimony (or rather, nontestimony) similar in almost all respects to that of Justices Ginsburg and Breyer.<sup>21</sup> This was so even though the Senate knew little about Justice Kennedy and still less about Justice

---

<sup>19</sup> See David A. Strauss, *Whose Confirmation Mess?*, *Am Prospect* 91, 96 (Summer 1994), reviewing Carter, *The Confirmation Mess*. Herein lies one of the mysteries of modern confirmation politics: given that the Republican Party has an ambitious judicial agenda and the Democratic Party has next to none, why is the former labeled the party of judicial restraint and the latter the party of judicial activism?

<sup>20</sup> *Id.* at 92, 95-96.

<sup>21</sup> Prior to nominating Justice Kennedy, the Reagan White House nominated Judge Douglas Ginsburg, only soon to withdraw the nomination. The decision to pull the nomination followed revelations about Judge Ginsburg's prior use of marijuana. Carter barely mentions this nomination. Carter, however, generally considers the prior illegal conduct of a nominee to be a meet subject for investigation, although not necessarily a sufficient reason for disqualification (pp 169-77).

Souter prior to the hearings—an ignorance which should have increased the importance of their testimony. (Just ask Senator Hatch whether he now wishes he had insisted that Justice Souter be more forthcoming.) The Senate also confirmed the nomination of Justice Thomas after his substantive testimony had become a national laughingstock. Take away the weakness of Justice Thomas's objective qualifications and the later charges of sexual harassment (inquiry into which Carter approves), and the Justice's Pinpoint, Georgia, testimonial strategy would have produced a solid victory.<sup>22</sup> This history offers scant support for Carter's lamentation that the confirmation process has become focused on a nominee's substantive testimony and obsessed with the nominee's likely voting record. So what, excepting once again Robert Bork, is Carter complaining about?

If Carter is right as to what makes a "confirmation mess," he had no reason to write this book—or at least to write it when he did. Senators today do not insist that any nominee reveal what kind of Justice she would make, by disclosing her views on important legal issues. Senators have not done so since the hearings on the nomination of Judge Bork. They instead engage in a peculiar ritual dance, in which they propound their own views on constitutional law, but neither hope nor expect the nominee to respond in like manner. Under Carter's criteria, this process ought to count as nothing more than a harmless charade, not as a problem of any real import. It is only if Carter's criteria are wrong—only if the hearings on Judge Bork ought to serve less as a warning than as a model—that we now may have a mess to clean up.

### III. CRITIQUING CARTER

What, then, of Carter's vision of the confirmation process? Should participants in the process accede to Carter's view of how to select a Supreme Court Justice? Or should they adopt a different, even an opposite, model?

One preliminary clarification is necessary. Carter's argument

---

<sup>22</sup> The margin of victory would have increased yet further had Thomas not made controversial statements, before his nomination, on subjects such as abortion and affirmative action. Carter is unclear as to whether (or how) participants in the confirmation process ought to take account of such prenomination statements. If Carter does approve of an evaluation of the substantive views expressed by a nominee in prior speeches or writings, then virtually all of the votes cast against Justice Thomas would have derived from the consideration of factors that Carter himself deems relevant to the process.

against a Bork-like confirmation process focuses entirely on the scope of the inquiry, not at all on the identity (executive or legislative) of the inquirer. This is an important point because other critics of the Bork hearings have rested their case on a distinction between the roles of the President and the Senate; they have argued that in assessing the substantive views of the nominee, the Senate ought to defer to the President.<sup>23</sup> Carter (I think rightly) rejects this claim, adopting instead the position that the Senate and the President have independent responsibility to evaluate, by whatever criteria are appropriate, whether a person ought to serve as a Supreme Court Justice.<sup>24</sup> Carter's argument concerns the criteria that the participants—that is, all the participants—in the confirmation process ought to use to make this decision. It is thus Carter's contention not merely that the Senate ought to forgo inquiry into a nominee's legal views and philosophy, but also that the President ought to do so—in short, that such inquiry, by whomever conducted, crosses the bounds of propriety. (And although Carter does not address the issue, his arguments apply almost equally well to an investigation of the views expressed in a person's written record as to an inquiry into the person's views by means of an oral examination.)

This analysis raises some obvious questions. If substantive inquiry is off-limits, on what basis will the President and Senate exercise their respective roles in the appointments process? Will this limited basis prove sufficient to evaluate and determine whether a nominee (or would-be nominee) should sit on the Court? Will an inquiry conducted on this basis appropriately educate and engage the public as to the Court's decisions and functions? Some closer exploration of Carter's views, as they relate to this set of issues, will illustrate at once the inadequacy of his proposals and the necessity for substantive inquiry of nominees, most notably in Senate hearings.

Carter argues that both the President and the Senate ought to pay close attention to a nominee's (or a prospective nominee's)

---

<sup>23</sup> See, for example, John O. McGinnis, *The President, the Senate, the Constitution, and the Confirmation Process: A Reply to Professors Strauss and Sunstein*, 71 *Tex L Rev* 633, 636, 653-54 (1993).

<sup>24</sup> This position has become common in the literature on the confirmation process. See David A. Strauss and Cass R. Sunstein, *The Senate, the Constitution, and the Confirmation Process*, 101 *Yale L J* 1491 (1992). See also Charles L. Black, Jr., *A Note on Senatorial Consideration of Supreme Court Nominees*, 79 *Yale L J* 657 (1970). Because Carter and I agree on the issue, and because the relevant arguments have been stated fully elsewhere, this Review addresses the issue only indirectly.

objective qualifications. There may be, as Carter notes, some disagreement as to what these are (pp 161-62). Must, for example (as Carter previously has argued<sup>25</sup>), a nominee have served on another appellate court—or may (as I believe) she demonstrate the requisite intelligence and legal ability through academic scholarship, the practice of law, or governmental service of some other kind? Carter writes that we must form a consensus on these issues and then rigorously apply it—so that the Senate, for example, could reject a nomination on the simple ground that the nominee lacks the qualifications to do the job (p 162). On this point, Carter surely is right. It is an embarrassment that the President and Senate do not always insist, as a threshold requirement, that a nominee's previous accomplishments evidence an ability not merely to handle but to master the "craft" aspects of being a judge. In this respect President Clinton's appointments stand as models. No one can say of his nominees, as no one ought to be able to say of any, that they lack the training, skills, and aptitude to do the work of a judge at the highest level.

But Carter cannot think—and on occasion reveals he does not think—that legal ability alone ought to govern, or as a practical matter could govern, either the President's or the Senate's decision. If there was once a time when we all could agree on the single "best" nominee—as, some say, all agreed on Cardozo—that time is long past, given the nature of the work the Supreme Court long has accomplished. As Carter himself concedes, most of the cases the Supreme Court hears require more than the application of "mundane and lawyerly" skills; these cases raise "questions requir[ing] judgment in the finding of answers, and in every exercise of interpretive judgment, there comes a crucial moment when the interpreter's own experience and values become the most important data" (p 151). Carter offers as examples flag burning, segregated schools, and executive power (p 151), and he could offer countless more; it should be no surprise by now that many of the votes a Supreme Court Justice casts have little to do with technical legal ability and much to do with conceptions of value. Imagine our response if President Clinton had announced that he had chosen his most recent nominee to the Supreme Court by conducting a lottery among Richard Posner, Stephen Breyer, and Laurence Tribe because they seemed to him the nation's three smartest lawyers. If we are all realists now, as the saying goes, it is in the sense that we understand a choice among

---

<sup>25</sup> See Stephen Carter, *The Confirmation Mess*, 101 Harv L Rev 1185, 1188 (1988).

these three to have large consequences and that we would view a lottery among them as demonstrating a deficient understanding of the judicial process.

Carter recommends, in light of the importance of a judge's values, that the President and Senate augment their inquiry into a person's legal ability with an investigation of the person's morality. He says that "[t]he issue, finally, is . . . what sort of person the nominee happens to be" (p 151); and he asks the President and Senate to determine whether a person "possess[es] the right moral instincts" by investigating whether her "personal moral decisions seem generally sound" (p 152). Here, too, it is easy to agree with Carter that this trait ought to play some role in the appointments process. Moral character, and the individual acts composing it, matter for two reasons (although Carter does not disentangle them). First, elevating a person who commits acts of personal misconduct (for example, sexual harassment) to the highest legal position in the nation sends all the wrong messages about the conduct that we as a society value and honor. Second, moral character, as Carter recognizes, sometimes will be "brought to bear on concrete cases," so that "the morally superior individual" may also "be the morally superior jurist," in the sense that her decisions will have a "salutary rather than destructive effect on the Court and the country" (p 153).

But focusing the confirmation process on moral character (even in conjunction with legal ability) would prove a terrible error. For one thing, such a focus would aggravate, rather than ease, the meanness that Carter rightly sees as marring the confirmation process (and, one might add, much of our politics). The "second" hearing on Clarence Thomas ought to have taught at least that lesson. When the subject is personal character, rather than legal principle, the probability, on all sides, of using gutter tactics exponentially increases. There are natural limits on the extent to which debate over legal positions can become vicious, hurtful, or sordid—but few on the extent to which discussion of personal conduct can descend to this level.

More important, an investigation of moral character will reveal very little about the values that matter most in the enterprise of judging. What makes the Richard Posner different from the Stephen Breyer different from the Laurence Tribe is not moral character or behavior, in the sense meant by Carter; I am reasonably sure that each of these persons is, in his personal life and according to Carter's standard, a morally exemplary individual. What causes them to differ as constitutional interpreters is

something if not completely, then at least partly, severable from personal morality: divergent understandings of the values embodied in the Constitution and the proper role of judges in giving effect to those values. Disagreement on these matters can cause (and has caused), among the most personally upright of judges, disagreement on every concrete question of constitutional law, including (or especially) the most important. It is therefore difficult to understand why we would make personal moral standards the focal point of a decision either to nominate or to confirm a person as a Supreme Court Justice.<sup>26</sup>

What must guide any such decision, stated most broadly, is a vision of the Court and an understanding of the way a nominee would influence its behavior. This vision largely consists of a view as to the kinds of decisions the Court should issue. The critical inquiry as to any individual similarly concerns the votes she would cast, the perspective she would add (or augment), and the direction in which she would move the institution.<sup>27</sup> I do not mean to say that the promotion of "craft values"—the building of a Court highly skilled in legal writing and reasoning and also finely attuned to pertinent theoretical issues—is at all unimportant. Justice Scalia by now has challenged and amused a decade's worth of law professors, which is no small thing if that is your profession; more seriously, the quality and intelligence (even if ultimate wrong-headedness) of much of Justice Scalia's work has instigated a debate that in the long run can only advance legal inquiry. But the bottom-line issue in the appointments process must concern the kinds of judicial decisions that will serve the country and, correlatively, the effect the nominee will have on the Court's decisions. If that is too results oriented

---

<sup>26</sup> It is also true that a person may engage in immoral behavior without allowing that immorality to influence his judicial decision making. Our government is replete with womanizers who always vote in sympathy with the goal of sexual equality; our Court has seen a former Ku Klux Klan member who well understood the constitutional evil of state-imposed racism. Perhaps the (im)moral conduct in these cases is all that matters; perhaps, in any event, we ought to rely on the (im)moral conduct as a solid, even if not a foolproof, indicator of future judicial behavior. But consideration of these cases may increase further our reluctance to make moral character the critical determinant of confirmation decisions.

<sup>27</sup> The President and Senate thus ought to evaluate the nominee (or potential nominee) in the context of the larger institution she would join if confirmed. They are not choosing a judge who will staff the Supreme Court alone; they are choosing a judge who will act and interact with eight other members. The qualities desirable in a nominee may take on a different cast when this fact is remembered. Most obviously, the benefits of diversity of viewpoint become visible only when the nominee is viewed as just one member of a larger body.

in Carter's schema, so be it—though even he notes that a critical question is whether the Court's decisions will have a "salutary" or a "destructive" impact on the country (p 153). It is indeed hard to know how to evaluate a governmental institution, or the individuals who compose it, except by the effect of their actions (or their refusals to take action) on the welfare of society.

If this is so, then the Senate's consideration of a nominee, and particularly the Senate's confirmation hearings, ought to focus on substantive issues; the Senate ought to view the hearings as an opportunity to gain knowledge and promote public understanding of what the nominee believes the Court should do and how she would affect its conduct. Like other kinds of legislative fact-finding, this inquiry serves both to educate members of the Senate and public and to enhance their ability to make reasoned choices. Open exploration of the nominee's substantive views, that is, enables senators and their constituents to engage in a focused discussion of constitutional values, to ascertain the values held by the nominee, and to evaluate whether the nominee possesses the values that the Supreme Court most urgently requires. These are the issues of greatest consequence surrounding any Supreme Court nomination (not the objective qualifications or personal morality of the nominee); and the process used in the Senate to serve the intertwined aims of education and evaluation ought to reflect what most greatly matters.<sup>28</sup> At least this is true in the absence of any compelling reasons, of prudence or propriety, to the contrary; later I will argue, as against Carter, that such reasons are nowhere evident.

The kind of inquiry that would contribute most to understanding and evaluating a nomination is the kind Carter would forbid: discussion first, of the nominee's broad judicial philosophy and, second, of her views on particular constitutional issues. By "judicial philosophy" (a phrase Carter berates without explanation), I mean such things as the judge's understanding of the role of courts in our society, of the nature of and values embodied in our Constitution, and of the proper tools and techniques of interpretation, both constitutional and statutory. A nominee's views on these matters could prove quite revealing: contrast, for example, how Antonin Scalia and Thurgood Marshall would have answered these queries, had either decided (which neither did) to

---

<sup>28</sup> To structure the process to avoid these issues would be akin to enacting a piece of legislation without trying to figure out or explain the legislation's principal consequences. I presume that no one would commend such an approach generally to Congress.

share his thoughts with the Senate. But responses to such questions can—and have—become platitudinous, especially given the interrogators' scant familiarity with jurisprudential matters.<sup>29</sup> And even when a nominee avoids this vice, her statements of judicial philosophy may be so abstract as to leave uncertain, especially to the public, much about their real-world consequences. Hence the second aspect of the inquiry: the insistence on seeing how theory works in practice by evoking a nominee's comments on particular issues—involving privacy rights, free speech, race and gender discrimination, and so forth—that the Court regularly faces. It is, after all, how the Court functions with respect to such issues that makes it, in Carter's words, either a "salutary" or a "destructive" institution.

A focus on substance in fact would cure some of the deficiencies in the confirmation process that Carter pinpoints. Carter says that the process turns "tiny ethical molehills into vast mountains of outrage" (p 8)—and he is right that we have seen these transformations. To note but one example, the amount of heat generated by a few senators (and the New York Times) concerning Justice Breyer's recusal practices far exceeded the significance of the issue. But this occurs precisely because we have left ourselves with nothing else to talk about. Rather than feeling able to confront directly the question whether Justice Breyer was too moderate, Senator Metzenbaum (and likewise the New York Times) fumed about an issue not nearly so important, either to them or to the public. Carter also says that participants in the process have attempted to paint nominees (particularly Judge Bork) as "radical monster[s]—far outside the mainstream of both morality and law" (p 127). But assuming, as seems true, that senators and others at times have engaged in distortion—it would be surprising if they hadn't—the marginalization of substantive inquiry that Carter favors only would encourage this practice. If evaluating (and perhaps rejecting) a nominee on the

---

<sup>29</sup> Carter often takes senators to task for failing to question nominees on constitutional theory with the appropriate level of sophistication and nuance. Although there is some truth to this criticism, it is mixed in Carter's account with a healthy measure of professorial condescension. Given the need to explain matters of constitutional theory to the public, at least a few senators do quite well. To the extent Carter's criticism has merit, the real problem is that senators now can expect answers only to high-blown questions of constitutional theory. Senators wander in the unfamiliar ground of constitutional theory because they cannot gain access to the real, and very familiar, world of decisions and consequences. See Robert F. Nagel, *Advice, Consent, and Influence*, 84 Nw U L Rev 858, 863 (1990) ("Senators are certainly qualified to consider the impact of the law's abstractions.").

basis of her substantive positions is appropriate only in the most exceptional cases, then the natural opponents of a nomination will have every incentive to—indeed, will need to—characterize the nominee as a “radical monster.” The way to promote reasoned debate thus lies not in submerging substantive issues, but in making them the centerpiece of the confirmation process.

Further, a commitment to address substantive issues need not especially disadvantage scholars and others who have left a “paper trail,” as the received wisdom intones and Carter accepts (p 38). The conventional view is that substantive inquiry promotes substantive ciphers; hence the hearings on Robert Bork led to the nomination of David Souter. But this occurs only because the cipher is allowed to remain so—only because substantive questioning is reserved for nominees who somehow have “opened the door” to it by once having committed a thought to paper. If questioning on substantive positions ever were to become the norm, the nominee lacking a publication record would have no automatic advantage over a highly prolific author. The success of a nomination in each case would depend on the nominee’s views, whether or not previously expressed in a law review or federal reporter. Indeed, a confirmation process devoted to substantive inquiry might favor nominees with a paper trail, all else being equal. If there was any reason for the Senate to have permitted the testimonial demurrals of Justices Breyer and Ginsburg, it was that their views already were widely known, in large part through scholarship and reported opinions—and that those views were widely perceived as falling within the appropriate range. When this is so, extended questioning on legal issues may seem hardly worth the time and effort.<sup>30</sup> More available writing thus might lead to less required testimony in a confirmation process committed to substantive inquiry.

Finally, a confirmation process focused on substantive views usually will not violate, in the way Carter claims, norms of judi-

---

<sup>30</sup> The value of questioning in such circumstances is almost purely educative; the inquiry is a means not of discovering what the nominee thinks, in order to decide whether confirmation is warranted, but instead of conveying to members of the public what the nominee thinks, in order to give them both an understanding of the Court and a sense of participating in its composition. This function is itself important, see text accompanying note 28; it may provide a reason for holding substantive hearings even when senators can make, and have made, a decision as to a nominee’s views prior to asking a single question (as senators could have and, for the most part, did about the views of Justices Breyer and Ginsburg). The need for such hearings, however, is much greater when (as was true for Justices Souter and Thomas) the prior record and writings of the nominee leave real uncertainty as to the nominee’s legal philosophy.

cial impartiality or independence. Carter's "blank slate" notion of impartiality of judgment—"appointing Justices who make up their minds about how to vote before they hear any arguments rather than after is a threat," fusses Carter (p 56)—is an especial red herring. Judges are not partial in deciding cases because they have strong opinions, or previously have expressed strong opinions, on issues involved in those cases. If they were, the Supreme Court would have to place, say, Justice Scalia in a permanent state of recusal, given that in the corpus of his judicial opinions he has stated unequivocal views on every subject of any importance. And the Senate would have had to reject, on this ground alone, the nomination of Justice Ginsburg, who not only had written about abortion rights<sup>31</sup>—perhaps the most contentious issue in contemporary constitutional law—but who testified in even stronger terms as to her current views on that issue.<sup>32</sup> That both suggestions are absurd indicates that we do not yet, thankfully enough, consider either the possession or the expression of views on legal issues—even when strongly held and stated—to be a judicial disqualification.

As for "judicial independence," Carter speaks as though the term were self-defining—and as though it meant that in appointing judges to a court, the President and Senate must refrain from considering what they will do once they arrive there. But this would be an odd kind of decision to leave in the hands of elected officials: far better, if such subjects were forbidden, to allow judges to name their own successors—or to cede the appointment power to some ABA committee. In fact, the placement of this decision in the political branches says something about its nature—says something, in particular, about its connection to the real-world consequences of judicial behavior. Indeed, contrary to Carter's view, the President and Senate themselves have a constitutional obligation to consider how an individual, as a judge, will read the Constitution: that is one part of what it means to preserve and protect the founding instrument. The value of judicial independence does not command otherwise, however much Carter tries to convert this concept into a thought-suppressing mantra. The judicial independence that we should focus on protecting resides primarily in the inability of political officials, once having placed a person on a court, to interfere with what she

---

<sup>31</sup> See, for example, Ruth Bader Ginsburg, *Some Thoughts on Autonomy and Equality in Relation to Roe v. Wade*, 63 NC L Rev 375 (1985).

<sup>32</sup> See, for example, Confirmation Hearings for Ginsburg at 268-69 (cited in note 5).

does there. That seems a fair amount of independence for any branch of government.

I do not mean to argue here that the President and Senate may ask, and a nominee (or potential nominee) must answer, any question whatsoever. Some kinds of questions, as Carter contends, do pose a threat to the integrity of the judiciary. Suppose, for example, that a senator asked a nominee to commit herself to voting a certain way on a case that the Court had accepted for argument. We would object—and we would be right to object—to this question, on the ground that any commitment of this kind, even though unenforceable, would place pressure on the judge (independent of the merits of the case) to rule in a certain manner. This would impede the judge's ability to make a free and considered decision in the case, as well as undermine the credibility of the decision in the eyes of litigants and the public. And once we accept the impermissibility of such a question, it seems we have to go still further. For there are ways of requesting and making commitments that manage to circumvent the language of pledge and promise, but that convey the same meaning; and these scantily veiled expressions pose dangers almost as grave as those of explicit commitments to the fairness, actual and perceived, of the judicial process.

But we do not have to proceed nearly so far down the road of silence as Carter and recent nominees would take us—to a place where comment of any kind on any issue that might bear in any way on any case that might at any time come before the Court is thought inappropriate.<sup>33</sup> There is a difference between a prohibition on making a commitment (whether explicit or implicit) and a prohibition on stating a current view as to a disputed legal question. The most recent drafters of the Model Code of Judicial Conduct acknowledged just this distinction when they adopted the former prohibition in place of the latter for candidates for judicial office.<sup>34</sup> Of course, there will be hard cases—cases in which reasonable people may disagree as to whether a nominee's statement of opinion manifests a settled intent to decide in a

---

<sup>33</sup> For a similar conclusion, see Steven Lubet, *Advice and Consent: Questions and Answers*, 84 Nw U L Rev 879 (1990).

<sup>34</sup> See pp 96-97. Compare Model Code of Judicial Conduct Canon 5(A)(3)(d) (1990), with Code of Judicial Conduct Canon 7(B)(1)(c) (1972). See generally *Buckley v Illinois Judicial Inquiry Board*, 997 F2d 224, 230 (7th Cir 1993) (Judge Posner noting the difference between these two kinds of prohibitions and holding the broader prohibition, on "announc[ing] . . . views on disputed legal or political issues," to violate the First Amendment).

particular manner a particular case likely to come before the Court. But many easy cases precede the hard ones: a nominee can say a great, great deal before making a statement that, under this standard, nears the improper. A nominee, as I have indicated before, usually can comment on judicial methodology, on prior caselaw, on hypothetical cases, on general issues like affirmative action or abortion. To make this more concrete, a nominee can do . . . well, what Robert Bork did. If Carter and recent nominees are right, Judge Bork's testimony violated many times a crucial norm of judicial conduct. In fact, it did no such thing; indeed it should serve as a model.

Return for a moment to those hearings, in which the Senate—and the American people—evaluated Robert Bork's fitness. Carter stresses the distortion, exaggeration, and vilification that occurred during the debate on the nomination. And surely these were present—most notably, as Carter notes, in the misdescription of Bork's opinion in *American Cyanimid*.<sup>35</sup> But the most striking aspect of the debate over the Bork nomination was not the depths to which it occasionally descended, but the heights that it repeatedly reached.<sup>36</sup> What Carter tongue-in-cheek calls "the famous national seminar on constitutional law" (p 6) was just that. The debate focused not on trivialities (Carter's "ethical molehills") but on essentials: the understanding of the Constitution that the nominee would carry with him to the Court. Senators addressed this complex subject with a degree of seriousness and care not usually present in legislative deliberation; the ratio of posturing and hyperbole to substantive discussion was much lower than that to which the American citizenry has become accustomed. And the debate captivated and involved that citizenry in a way that, given the often arcane nature of the subject matter, could not have been predicted. Constitutional law became, for that brief moment, not a project reserved for judges, but an enterprise to which the general public turned its attention and contributed.

Granted that not all subsequent confirmation hearings could, or even should, follow the pattern set by the Bork hearings, in either their supercharged intensity or their attention to substance. A necessary condition of both was the extreme conservatism of Bork's known views, which made him an object of terror to some

---

<sup>35</sup> *Oil, Chemical & Atomic Workers Intl. Union v American Cyanimid Co.*, 741 F2d 444 (DC Cir 1984).

<sup>36</sup> For a similar view, see Strauss, *Am Prospect* at 94 (cited in note 19).

senators and veneration to others. It would be difficult to imagine hearings of the same kind following the nomination of Justice Ginsburg or Justice Breyer—two well-known moderates whose nominations had been proposed by senators on both sides of the aisle. To insist that these hearings take the identical form as the hearings on Judge Bork is not only to blink at political reality, but also to ignore the very real differences in the nature of the nominations.

But that said, the real “confirmation mess” is the gap that has opened between the Bork hearings and all others (not only for Justices Ginsburg and Breyer, but also, and perhaps especially, for Justices Kennedy, Souter, and Thomas). It is the degree to which the Senate has strayed from the Bork model. The Bork hearings presented to the public a serious discussion of the meaning of the Constitution, the role of the Court, and the views of the nominee; that discussion at once educated the public and allowed it to determine whether the nominee would move the Court in the proper direction. Subsequent hearings have presented to the public a vapid and hollow charade, in which repetition of platitudes has replaced discussion of viewpoints and personal anecdotes have supplanted legal analysis. Such hearings serve little educative function, except perhaps to reinforce lessons of cynicism that citizens often glean from government. Neither can such hearings contribute toward an evaluation of the Court and a determination whether the nominee would make it a better or worse institution. A process so empty may seem ever so tidy—muted, polite, and restrained—but all that good order comes at great cost.

And what is worse even than the hearings themselves is a necessary condition of them: the evident belief of many senators that serious substantive inquiry of nominees is usually not only inessential, but illegitimate—that their insistent questioning of Judge Bork was justified, if at all, by his overt “radicalism” and that a similar insistence with respect to other nominees, not so obviously “outside the mainstream,” would be improper. This belief is not so often or so clearly stated; but it underlies all that the Judiciary Committee now does with respect to Supreme Court nominations. It is one reason that senators accede to the evasive answers they now have received from five consecutive nominees. It is one reason that senators emphasize, even in posing questions, that they are asking the nominee only about philosophy and not at all about cases—in effect, inviting the nominee to spout legal theory, but to spurn any demonstration of

what that theory might mean in practice. It is one reason that senators often act as if their inquiry were a presumption—as if they, mere politicians, have no right to ask a real lawyer (let alone a real judge) about what the law should look like and how it should work. What has happened is that the Senate has absorbed criticisms like Carter's and, in so doing, has let slip the fundamental lesson of the Bork hearings: the essential rightness—the legitimacy and the desirability—of exploring a Supreme Court nominee's set of constitutional views and commitments.

The real confirmation mess, in short, is the absence of the mess that Carter describes. The problem is not that the Bork hearings have set a pattern for all others; the problem is that they have not. And the problem is not that senators engage in substantive discussion with Supreme Court nominees; the problem is that they do not. Senators effectively have accepted the limits on inquiry Carter proposes; the challenge now is to overthrow them.

In some sense, Carter is right that we will clean up the mess only when we change "our attitudes toward the Court as an institution"—when we change the way we "view the Court" (p 188). But as he misdescribes the mess, so too does Carter misapprehend the needed attitudinal adjustment. We should not persuade ourselves, as Carter urges, to view the Court as a "mundane and lawyerly" institution and to view the position of Justice as "simply a job" (pp 205-06). We must instead remind ourselves to view the Court as the profoundly important governmental institution that, for good or for ill, it has become and, correlatively, to view the position of Justice as both a seat of power and a public trust. It is from this realistic, rather than Carter's nostalgic, vision of the Court that sensible reform of the confirmation process one day will come. And such reform, far from blurring a nominee's judicial philosophy and views, will bring them into greater focus.

VOLUME 60 NUMBERS 3 & 4  
SUMMER/FALL 1993

# The University of Chicago Law Review

Words, Conduct, Caste

*Cass R. Sunstein*

Girls Should Bring Lawsuits Everywhere . . .  
Nothing Will be Corrupted:  
Pornography as Speech and Product  
Regulation of Hate Speech and Pornography  
After *R.A.V.*

*Marianne Wesson*

*Elena Kagan*

A Systematic Approach to Privilege Against  
Self-Incrimination Claims When  
Foreign Prosecution Is Feared

The Appealability of Class Action Settlements  
by Unnamed Parties

Are Viatical Settlements Securities Within  
the Regulatory Control of the Securities  
Act of 1933?

Harmless Error as Constitutional Common Law:  
Congress's Power to Reverse *Arizona v*  
*Fulminante*

*Chevron* and the Canon Favoring Indians

Rule 11 and Federal Sovereign Immunity:  
Respecting the Explicit Waiver Requirement

The Bricoleur at the Center

*The Partial Constitution*

Cass R. Sunstein

*Mark Tushnet*

## Regulation of Hate Speech and Pornography After *R.A.V.*

Elena Kagan†

This Essay on the regulation of hate speech and pornography addresses both practicalities and principles. I take it as a given that we live in a society marred by racial and gender inequality, that certain forms of speech perpetuate and promote this inequality, and that the uncoerced disappearance of such speech would be cause for great elation. I do not take it as a given that all governmental efforts to regulate such speech thus accord with the Constitution. What is more (and perhaps what is more important), the Supreme Court does not, and will not in the foreseeable future, take this latter proposition as a given either. If confirmation of this point were needed, it came last year in the shape of the Court's opinion in *R.A.V. v City of St. Paul*.<sup>1</sup> There, the Court struck down a so-called hate speech ordinance, in the process reiterating, in yet strengthened form, the tenet that the First Amendment presumptively prohibits the regulation of speech based upon its content, and especially upon its viewpoint. That decision demands a change in the nature of the debate on pornography and hate speech regulation. It does so for principled reasons—because it raises important and valid questions about which approaches to the regulation of hate speech and pornography properly should succeed in the courts. And it does so for purely pragmatic reasons—because it makes clear that certain approaches almost surely will not succeed.

In making this claim, I do not mean to suggest that all efforts to regulate pornography and hate speech be suspended, on the

---

† Assistant Professor of Law, The University of Chicago Law School. This Essay is based on remarks I made in a panel discussion at the conference, "Speech, Equality, and Harm: Feminist Legal Perspectives on Pornography and Hate Propaganda," held at The University of Chicago Law School on March 5-7, 1993. The argument has been expanded only slightly, and the reader is asked to make allowances for my necessarily abbreviated discussion of many complicated issues. I am grateful to Mary Becker, Larry Lessig, Michael McConnell, Geoffrey Stone, David Strauss, and Cass Sunstein for valuable advice and comments. The Class of '64 Fund and the Russell Parsons Faculty Research Fund at The University of Chicago Law School provided financial support.

<sup>1</sup> 112 S Ct 2538 (1992).

ground either of mistake or of futility. Quite the opposite. *R.A.V.* largely forecloses some lines of advocacy and argument (until now the dominant lines), as well perhaps it should have. But the decision leaves open alternative means of regulating some pornography and hate speech, or of alleviating the harms that such speech causes. The primary purpose of this Essay is to offer some of these potential new approaches for consideration and debate. The question I pose is whether there are ways to achieve at least some of the goals of the anti-pornography and anti-hate speech movements without encroaching on valuable and ever more firmly settled First Amendment principles. This Essay is just that—an essay, a series of trial balloons, which may be shot down, from either side or no side at all, by me or by others. The point throughout is to emphasize the range of approaches remaining available after *R.A.V.* and meriting discussion.

### I. THE PROBLEM OF VIEWPOINT DISCRIMINATION

In *R.A.V.*, the Court struck down a local ordinance construed to prohibit those fighting words, but only those fighting words, based on race, color, creed, religion, or gender.<sup>2</sup> Fighting words long have been considered unprotected expression—so valueless and so harmful that government may prohibit them entirely without abridging the First Amendment.<sup>3</sup> Why, then, was the ordinance before the Court constitutionally invalid? The majority reasoned that the ordinance's fatal flaw lay in its incorporation of a kind of content-based distinction. The ordinance, on its very face, distinguished among fighting words on the basis of their subject matter: only fighting words concerning "race, color, creed, religion or gender" were forbidden.<sup>4</sup> More, and much more nefariously in the Court's view, the ordinance in practice discriminated between different viewpoints: it effectively prohibited racist and sexist fighting words, while allowing all others.<sup>5</sup> Antipathy to such viewpoint distinctions, the Court stated, lies at the heart of the guarantee of freedom of expression. "The government may not regulate [speech] based on hostility—or favoritism—towards the underlying

---

<sup>2</sup> Id at 2542. The Supreme Court defined "fighting words" in *Chaplinsky v New Hampshire*, 315 US 568, 572 (1942), as words "which by their very utterance inflict injury or tend to incite an immediate breach of the peace."

<sup>3</sup> *Chaplinsky*, 315 US at 572.

<sup>4</sup> *R.A.V.*, 112 S Ct at 2541, 2547.

<sup>5</sup> Id at 2547-48.

message expressed"; it may not suppress or handicap "particular ideas."<sup>6</sup>

The reasoning in *R.A.V.* closely resembles that found in the key judicial decision on the regulation of pornography. In *American Booksellers Ass'n, Inc. v Hudnut*,<sup>7</sup> affirmed summarily by the Supreme Court, the United States Court of Appeals for the Seventh Circuit invalidated the Indianapolis anti-pornography ordinance drafted by Andrea Dworkin and Catharine MacKinnon. That ordinance declared pornography a form of sex discrimination, with pornography defined as "the graphic sexually explicit subordination of women, whether in pictures or in words," that depicted women in specified sexually subservient postures.<sup>8</sup> The core problem for the Seventh Circuit, as for the Supreme Court in *R.A.V.*, was one of viewpoint discrimination. The ordinance, according to the Court of Appeals, made the legality of expression "depend[ent] on the perspective the author adopts."<sup>9</sup> Sexually explicit speech portraying women as equal was lawful; sexually explicit speech portraying women as subordinate was not. The ordinance, in other words, "establishe[d] an 'approved' view" of women and of sexual relations.<sup>10</sup> From this feature, invalidation necessarily followed: "The state may not ordain preferred viewpoints in this way. The Constitution forbids the state to declare one perspective right and silence opponents."<sup>11</sup>

The approach used in *R.A.V.* and *Hudnut* has a large body of case law behind it. The presumption against viewpoint discrimination did not emerge alongside of, or in response to, the effort to curtail certain forms of racist and sexist expression. Rather, that presumption long has occupied a central position in First Amendment doctrine. Decades ago, for example, the Supreme Court employed the presumption to strike down laws restricting expression that discredited the military or that presented adultery in a favorable light, and more recently, the Court invoked the presumption to invalidate flag-burning statutes.<sup>12</sup> This is not to say that the Court invariably has invalidated laws that incorporate view-

---

<sup>6</sup> Id at 2545, 2549.

<sup>7</sup> 771 F2d 323 (7th Cir 1985), aff'd mem, 475 US 1001 (1986).

<sup>8</sup> Id at 324.

<sup>9</sup> Id at 328.

<sup>10</sup> Id.

<sup>11</sup> Id at 325.

<sup>12</sup> See *Schacht v United States*, 398 US 58, 67 (1970) (military); *Kingsley Int'l Pictures Corp. v Regents*, 360 US 684, 688 (1959) (adultery); *Texas v Johnson*, 491 US 397, 416-17 (1989) (flag-burning); *United States v Eichman*, 496 US 310, 317-18 (1990) (same).

point favoritism. Exceptions to the rule exist, although the Court rarely has seen fit to acknowledge them as such; in a number of areas of First Amendment law (and especially when so-called low-value speech is implicated), the Court breezily has ignored both more and less obvious forms of viewpoint preference.<sup>13</sup> Still, the rule has been more often honored than honored in the breach, and the Supreme Court's opinion in *R.A.V.*, as well as its summary affirmation of *Hudnut*, could have been expected.

Moreover, the Court's decision in *R.A.V.* entrenched still further the presumption against viewpoint-based regulation of speech. To be sure, the majority opinion received only five votes and came under vehement attack from the remaining Justices.<sup>14</sup> Thus, some might reason that the disposition of the case reveals a weakening in the Court's commitment to viewpoint neutrality, either across the board or with respect to racist and sexist expression. If this reasoning were valid, those disliking *R.A.V.* might simply wait and pray for an advantageous change in the Court's membership. But any such reading of the case rests on a grave misunderstanding. The Court's opinion received the support of only a bare majority because, for two reasons having nothing to do with the particular viewpoint involved, the case appeared to some Justices not to invoke the presumption against viewpoint regulation at all. First,

---

<sup>13</sup> Several examples of this blindness to viewpoint discrimination occur in the area of commercial speech. See *Posadas de Puerto Rico Associates v. Tourism Co. of Puerto Rico*, 478 US 328, 330-31 (1986) (upholding a law prohibiting advertising of casino gambling, but leaving untouched all speech discouraging such gambling); *Central Hudson Gas & Electric v. Public Service Commission*, 447 US 557, 569-71 (1980) (striking down a broad law prohibiting advertising to stimulate the use of electricity, but suggesting that a more narrowly-tailored law along the same lines would meet constitutional standards, even if the law were to allow all expression discouraging use of electricity). In addition, as Catharine MacKinnon has noted, the delineation of entire low-value categories of speech, such as obscenity and child pornography, may be thought to reflect a kind of viewpoint discrimination, given that the speech falling within such categories likely expresses a single (disfavored) viewpoint about sexual matters. See Catharine A. MacKinnon, *Feminism Unmodified: Discourses on Life and Law* 212 (Harvard, 1987). Further discussion of this point, and its relevance for the regulation of pornography and hate speech, appears in note 73 and the text accompanying note 80. Finally, the Court has indicated that the usual presumption against viewpoint discrimination does not apply, or at least does not apply in full force, when the government engages in selective funding of speech, rather than selective restriction of speech. See *Rust v. Sullivan*, 111 S. Ct. 1759, 1772-73 (1991); text accompanying notes 28-29.

<sup>14</sup> The four Justices who refused to join the Court's opinion also voted to invalidate the St. Paul ordinance, but only because of a concern about overbreadth that easily could have been corrected. They assailed the majority's conclusion that the presumption against viewpoint discrimination mandated invalidation of the statute, either on the view that the presumption failed to operate in spheres of unprotected speech, see 112 S. Ct. at 2551-54 (White concurring) and *id.* at 2560 (Blackmun concurring), or on the view that the ordinance incorporated no viewpoint-based distinction, see *id.* at 2570-71 (Stevens concurring).

and most important, the alleged viewpoint discrimination in the case occurred within a category of speech—fighting words—that the Court long ago declared constitutionally unprotected. Second, the viewpoint discrimination found in the ordinance existed not on its face, but only in application—and even in application, only with a fair bit of argument.<sup>15</sup> Had the law distinguished on its face between racist (or sexist) speech and other speech outside the category of fighting words, the Court's decision likely would have been unanimous.<sup>16</sup> What *R.A.V.* shows, then, is the depth, not the tenuousness, of the Court's commitment to a viewpoint neutrality principle. And what *R.A.V.* did, in applying that principle to a case of non-facial discrimination in an unprotected sphere, was to render that principle even stronger.

Any attempt to regulate pornography or hate speech—or at least any attempt standing a chance of success—must take into account these facts (the “is,” regardless whether the “ought”) of First Amendment doctrine. A law specifically disfavoring racist or sexist speech (or, to use another construction, a law distinguishing between depictions of group members as equal and depictions of group members as subordinate) runs headlong into the longstanding, and newly revived, principle of viewpoint neutrality. I do not claim that exceptions to this principle will never be made, or even that such exceptions will not be made by the current Court. Exceptions, as noted previously, have been recognized before (even if not explicitly); they doubtless will be recognized again; and in the last section of this Essay, I consider briefly whether and how to frame them. I do claim that given the current strength of the viewpoint neutrality principle, a purely pragmatic approach to regulating hate speech and pornography would seek to use laws not subject to the viewpoint discrimination objection, while also seeking to justify—as exceptions—carefully crafted and limited departures from the rule against viewpoint regulation.

---

<sup>15</sup> The St. Paul ordinance, on its face, discriminated only on the basis of subject matter, as the Court conceded. For the dispute on whether the ordinance applied in a viewpoint-discriminatory manner, contrast the majority opinion, 112 S Ct at 2547-48, with the concurring opinion of Justice Stevens, *id.* at 2570-71. Contrast also Cass R. Sunstein, *On Analogical Reasoning*, 106 Harv L Rev 741, 762-63 & n 78 (1993) (*R.A.V.* ordinance not viewpoint-based in practice), with Elena Kagan, *The Changing Faces of First Amendment Neutrality: R.A.V. v St. Paul, Rust v Sullivan, and the Problem of Content-Based Underinclusion*, 1992 S Ct Rev 29, 69-71 (*R.A.V.* ordinance viewpoint-based in practice).

<sup>16</sup> See note 14 for a description of the concurring Justices' objections to the Court's decision. In the case hypothesized in the text, those objections would have evaporated.

This approach, in my view, also best accords with important free speech principles (the "ought" in the "is" of First Amendment doctrine). A focus on the feasible is arguably irresponsible if the feasible falls desperately short of the proper. But here, I think, that is not the case. If reality—the current state of First Amendment doctrine—counsels certain proposals and not others, certain lines of argument and not others, so too do important values embodied in that doctrine. More specifically, the principle of viewpoint neutrality, which now stands as the primary barrier to certain modes of regulating pornography and hate speech, has at its core much good sense and reason. Although here I can do no more than touch on the issue, my view is that efforts to regulate pornography and hate speech not only will fail, but also should fail to the extent that they trivialize or subvert this principle.

Those who have criticized the courts for using the viewpoint neutrality principle against efforts to regulate pornography or hate speech usually have offered one of two arguments. First, some have claimed that such efforts comport with the norm of viewpoint neutrality because they are based on the harm the speech causes, rather than the viewpoint it espouses.<sup>17</sup> Second, and more dramatically, some have challenged the norm itself as incoherent, worthless, or dangerous.<sup>18</sup> Both lines of argument have enriched discussion of the viewpoint neutrality principle, by challenging the tendency of such discussion to do nothing more than apotheosize. Yet both approaches, in somewhat different ways, slight the reasons and values underlying current First Amendment doctrine—including the decisions in *R.A.V.* and *Hudnut*.

The claim that pornography and hate-speech regulation is harm-based, rather than viewpoint-based, has an initial appeal, but turns out to raise many hard questions. The claim appeals precisely because it reflects an understanding of the value of a view-

---

<sup>17</sup> See, for example, Cass R. Sunstein, *Pornography and the First Amendment*, 1986 Duke L J 589, 612; MacKinnon, *Feminism Unmodified* at 212 (cited in note 13). See also *R.A.V.*, 112 S Ct at 2570 (Stevens concurring). Professor Sunstein always has combined this argument with a fuller analysis of when exceptions to the viewpoint regulation doctrine are justified; for him, the ability to classify a law as harm-based seems not the end, but only the start of the inquiry. See Cass R. Sunstein, *Words, Conduct, Caste*, 60 U Chi L Rev 795, 796 (1993) (in this issue). My brief discussion, in Section II of this Essay, on whether and when to recognize such exceptions owes much to his work on the subject.

<sup>18</sup> See Mary Becker, *Conservative Free Speech and the Uneasy Case for Judicial Review*, 64 U Colo L Rev 975, 1044-47 (1993) (arguing that a viewpoint neutrality norm harms women and minority groups); MacKinnon, *Feminism Unmodified* at 210-13 (cited in note 13) (challenging the ability to identify viewpoint regulation except by reference to social consensus).

point neutrality norm and a desire to maintain it: if pornography and hate-speech regulation is harm-based, then we can have both it and a rule against viewpoint discrimination.<sup>19</sup> But the two yearnings may not be so easy to accommodate, for it is not clear that the classification proposed can support much weight. It is true that statutory language can focus either on the viewpoint of speech or on the injury it causes: contrast an ordinance that prohibits "sexually explicit materials approving the subordination of women" with an ordinance that prohibits "sexually explicit materials causing the subordination of women."<sup>20</sup> But if we assume (as a meaningful system of free speech must) that speech has effects—that the expression of a view will often cause people to act on it—then the two phrasings should be considered identical for First Amendment purposes. To grasp this point, consider here a few further examples. Contrast a law that prohibits criticism of the draft with a law that prohibits any speech that might cause persons to resist the draft.<sup>21</sup> Or, to use a case with more contemporary resonance, contrast an ordinance punishing abortion advocacy and counseling with an ordinance punishing any speech that might induce a woman to get an abortion. To sever these pairs of statutes would be to transform the First Amendment into a formal rule of legislative drafting, concerned only with appearance. In all these cases, the facially harm-based statute and the facially viewpoint-based statute function in the same way, because it is speech of a certain viewpoint, and only of that viewpoint, which causes the alleged injury. The facially harm-based statute in these circumstances will curtail expression of a particular message as surely as will the statute that refers to the message in explicit language. Given this functional identity, the statutes properly are viewed as cognates.<sup>22</sup>

---

<sup>19</sup> I suspect that a wish of this kind explains Justice Stevens's insistence in *R.A.V.* that the St. Paul ordinance regulated speech "not on the basis of . . . the viewpoint expressed, but rather on the basis of the *harm* the speech causes." 112 S Ct at 2570 (Stevens concurring). Both in *R.A.V.* and in numerous other opinions and articles, Justice Stevens has expressed unwavering support for the presumption against viewpoint regulation. For the most recent example, see The Hon. John Paul Stevens, *The Freedom of Speech*, 102 Yale L J 1293, 1309 (1993).

<sup>20</sup> The example, in slightly different form, appears in Geoffrey R. Stone, *Anti-Pornography Legislation as Viewpoint-Discrimination*, 9 Harv J L & Pub Pol 461, 467 (1986). As Stone points out, the MacKinnon-Dworkin ordinance, as written, is at any rate closer to the law focusing on the viewpoint espoused than to the law focusing on the harm caused. *Id.*

<sup>21</sup> This example also appears in Stone. *Id.*

<sup>22</sup> An argument to the contrary might rely not on the effects of the statutes, but on the intent of the legislature in passing them. The claim here would be that the facially harm-based statute more likely springs from a legitimate governmental motive than does the facially viewpoint-based statute. But this claim seems dubious in any case in which the

This equivalence does not by itself destroy the claim that pornography regulation is harm-based, because both versions of the law might be characterized in this manner: so long as a legislature reasonably decides, as it surely could with respect to pornography, that speech causes harm, then regulation responding to that harm (however framed) might be considered neutral, rather than an effort to disfavor certain viewpoints. But this approach, too, makes any distinction between viewpoint-based regulation and harm-based regulation collapse upon itself. Using this analysis, almost all viewpoint-based regulation can be described as harm-based, responding neutrally not to ideas as such, but to their practical consequences. For it is difficult to see why anyone would opt to regulate a viewpoint that did not cause what seemed (to the regulators at least) to be a harm—or at a bare minimum, that could not reasonably be described as harmful. So, to return to the examples used above, a law prohibiting criticism of the draft could be termed harm-based given that such speech in fact produces draft resistance; or a law prohibiting abortion counseling and advocacy could be termed harm-based given that such speech in fact increases the incidence of abortion (which many would count a serious injury). The substitution of labels—“harm-based” for “viewpoint-based”—thus either allows most viewpoint regulation to go forward or leaves yet unanswered the central issue of precisely when such regulation is appropriate.

The more extreme critique of a case like *Hudnut*—that viewpoint discrimination doctrine is both incoherent and corrupt—is in many ways more difficult to counter. This critique rebels against the very core of First Amendment doctrine by accepting the government’s power to suppress viewpoints as such whenever the viewpoints are thought to cause some requisite harm.<sup>23</sup> But the justification for this position includes at least one extremely potent point: that recognizing viewpoint regulation may well depend on the decisionmaker’s viewpoint; more specifically, that a judicial

---

statutes in fact operate in a similar manner. Because the legislators will know that the facially harm-based statute, like the facially viewpoint-based statute, will succeed in curtailing a specific message, their decision to phrase the statute in terms of harm (especially in light of a legal rule that effectively counsels them to do so) cannot provide a guarantee of legitimate intent.

<sup>23</sup> See MacKinnon, *Feminism Unmodified* at 212-13 (cited in note 13). Even under current First Amendment doctrine, the government may engage in viewpoint discrimination in emergency circumstances amounting to something like a clear and present danger. The critique discussed in the text would allow viewpoint regulation on a much less stringent showing.

decisionmaker will be least likely to recognize (or count as relevant) viewpoint regulation when the regulator's viewpoint lines up with his own.<sup>24</sup> This phenomenon may explain in part the willingness of courts to accept anti-obscenity laws at the same time as they strike down anti-pornography laws.<sup>25</sup> More generally, this epistemological problem may skew viewpoint discrimination doctrine, as it operates in practice, in favor of the status quo—resulting in the disproportionate approval of laws most reflective of traditional sentiment and the disproportionate invalidation of laws least so.

But even assuming this is true, I doubt that the appropriate response lies in undermining, let alone eliminating, the viewpoint discrimination principle. That principle grows out of two concerns, as meaningful today as ever in the past.<sup>26</sup> The first relates to the effects of viewpoint discrimination: such action skews public debate on an issue by restricting the ability of one side (and one side only) to communicate a message. The second relates to governmental purposes: viewpoint regulation often arises from hostility toward ideas as such, and this disapproval constitutes an illegitimate justification for governmental action. Of course, particular instances of viewpoint discrimination may spring from benign purposes and have benign effects. Legislators may engage in viewpoint discrimination in an effort not to suppress ideas, but to respond to real harms; and the resulting damage to public discourse may signify little when measured against the harms averted. But how are the courts, or the people, or even legislators themselves to make these determinations of motive and effect in any given case? Will it not always be true that a benign motive can be assigned to governmental action? Will not any judgment as to relative harms depend on an evaluation of the message affected? From these questions, relating to the difficulty of evaluating particular purposes and effects, emerges a kind of rule-utilitarian justification for the ban on viewpoint discrimination.

The historic examples of the dangers of viewpoint discrimination, on the counts of both purpose and effect, are well-known and legion: the government's attempts, especially during World War I, to stifle criticism of military activities; its attempts in the 1950s to

---

<sup>24</sup> See *id.* at 212; Becker, 64 U Colo L Rev at 1046-47 (cited in note 18).

<sup>25</sup> For discussion of the viewpoint bias inherent in obscenity laws, see notes 13 and 73 and text accompanying note 80.

<sup>26</sup> The classic discussion of the bases for viewpoint discrimination doctrine is Geoffrey R. Stone, *Content Regulation and the First Amendment*, 25 Wm & Mary L Rev 189 (1983).

suppress support of Communism; its efforts, stretching over decades, to prevent the burning of American flags as a means of protesting the government and its policies.<sup>27</sup> And if all these seem remote either from current threats or from the kind of viewpoint regulation at issue in *Hudnut* and *R.A.V.*—if they seem the stories of another generation, with little relevance for today—consider instead the case of *Rust v Sullivan*,<sup>28</sup> previewed in earlier hypotheticals. There, the government favored anti-abortion speech over abortion advocacy, counseling, and referral, and the Court, to its discredit, announced that because the selectivity occurred in the context of a governmental funding program, the presumption against viewpoint discrimination was suspended.<sup>29</sup> Or instead consider the numerous ways in which some of the strange bedfellows of anti-pornography feminists (and one must admit their presence) might choose (indeed, have chosen) to attack the expression of, among others, gays and lesbians.

The key point here is only strengthened by the insight that viewpoint discrimination doctrine, as applied by the courts, has a way of producing some patterned inconsistencies; or to put this another way, the very critique of the Court's viewpoint discrimination doctrine exposes the need for a viewpoint neutrality principle. For what the critique highlights is the tendency of governmental actors (of all kinds) to see speech regulation through the lens of their own orthodoxies, as well as the ease with which such orthodoxies can thereby become entrenched. Recognition of this process lies at the very core of the viewpoint discrimination doctrine: as Justice Stevens recently has noted, that doctrine responds, preeminently, to fear of the "imposition of an official orthodoxy,"<sup>30</sup> even (or perhaps especially) as to matters involving sex or race. That judicial decisionmakers, in applying the doctrine, sometimes will succumb to the views they hold hardly argues in favor of granting carte blanche to legislative decisionmakers to bow to theirs. It is difficult to see how women and minorities, who have the most to lose from the establishment of political orthodoxy,

---

<sup>27</sup> See Akhil Reed Amar, *The Supreme Court, 1991 Term—Comment: The Case of the Missing Amendments: R.A.V. v. City of St. Paul*, 106 Harv L Rev 124 (1992), for a comparison of *R.A.V.* and the Court's most recent flag-burning cases, *Texas v Johnson*, 491 US 397 (1989), and *United States v Eichman*, 496 US 310 (1990).

<sup>28</sup> 111 S Ct 1759 (1991).

<sup>29</sup> *Id.* at 1771-73. For a comparison of *Rust* and *R.A.V.*, see Kagan, 1992 S Ct Rev 29 (cited in note 15).

<sup>30</sup> Stevens, 102 Yale L J at 1304 (cited in note 19).

would gain by jettisoning the First Amendment doctrine that most protects against this prospect.

None of this discussion, of course, denies either the possibility or the desirability of crafting carefully circumscribed exceptions to First Amendment norms of viewpoint neutrality, and in the last section of this Essay, I briefly consider whether and how this task might be accomplished. Perhaps more important, none of this discussion gainsays the possibility of responding to the harms of pornography and hate speech through measures that do not contravene these norms. It is surely these measures, viewed from a pragmatic perspective, that stand the best chance of succeeding. And it usually will be these measures that pose the least danger to free speech principles. I turn, then, to a consideration of such proposals, less with the aim of making specific recommendations than with the aim of injecting new questions into the debate on hate speech and pornography regulation.

## II. NEW APPROACHES

I canvass here four general approaches; each is capable of encompassing many specific proposals. The four approaches are, in order: (1) the enactment of new, or the stricter use of existing, bans on conduct; (2) the enactment of certain kinds of viewpoint-neutral speech restrictions; (3) the enhanced use of the constitutionally unprotected category of obscenity; and (4) the creation of carefully supported and limited exceptions to the general rule against viewpoint discrimination. The proposals I outline within these approaches are meant to be illustrative, rather than exhaustive. Many fall well within constitutional boundaries; others test (or, with respect to the fourth approach, directly challenge) the current parameters. The latter proposals raise hard questions relating to whether they (no less than the standard viewpoint-based regulation) too greatly subvert principles necessary to a system of free expression. I will touch on many of these questions, although I cannot give them the extended treatment they merit.

### A. Conduct

The most obvious way to avoid First Amendment requirements is to regulate not speech, but conduct. Recently, some schol-

ars have sought to meld these two together.<sup>31</sup> Speech is conduct, they say, because speech has consequences (speech, that is, "does" something); or conduct is speech because conduct has roots in ideas (conduct, that is, "says" something). I use these terms in a different sense. When "conduct" becomes a synonym for "speech" (or "speech" for "conduct"), the command of the First Amendment becomes incoherent; depending on whether the paradigm of conduct or speech holds sway, government can regulate either almost everything or almost nothing. The speech/conduct line is hard to draw, but it retains much meaning in theory, and even more in practice. When I say "conduct," then, I mean acts that, in purpose and function, are not primarily expressive.<sup>32</sup> The government can regulate such acts without running afoul of the First Amendment.<sup>33</sup> Here, I discuss two specific kinds of conduct regulation: the continued enactment and use of hate crimes laws and the increased application of legal sanctions for acts commonly performed in the making of pornography.

The typical hate crimes law, as the Supreme Court unanimously ruled last Term, presents no First Amendment problem.<sup>34</sup> Hate crimes laws, as usually written, provide for the enhancement of criminal penalties when a specified crime (say, assault) is committed because of the target's race, religion, or other listed status.<sup>35</sup> These laws are best understood as targeting not speech, but acts—because they apply regardless whether the discriminatory conduct at issue expresses, or is meant to express, any sort of message. In this way, hate crimes laws function precisely as do other discrimination laws—for example, in the sphere of employment.<sup>36</sup>

---

<sup>31</sup> See MacKinnon, *Feminism Unmodified* at 129-30, 193-94 (cited in note 13); Charles R. Lawrence III, *If He Hollers Let Him Go: Regulating Racist Speech on Campus*, 1990 Duke L J 431, 438-44.

<sup>32</sup> The approach, in focusing on expressive quality, is similar to the analysis that Cass Sunstein presents in these pages. See *Words, Conduct, Caste*, 60 U Chi L Rev at 807-09 (cited in note 17). See also Amar, 106 Harv L Rev at 133-39 (cited in note 27). Of course, as sketched here, the definition begs all kinds of questions about when acts, either in purpose or in function, are primarily expressive.

<sup>33</sup> So, for example, it goes without saying that the City of St. Paul could have proceeded against the juvenile offenders in *R.A.V.* through the law of trespass. See *R.A.V.*, 112 S Ct at 2541 n 1 (listing other statutes under which the offenders could have been punished).

<sup>34</sup> *Wisconsin v Mitchell*, 113 S Ct 2194 (1993).

<sup>35</sup> See, for example, Cal Penal Code § 422.7 (West 1988 & Supp 1993); NY Penal Law § 240.30(3) (McKinney Supp 1993); Or Rev Stat § 166.165(1)(a)(A) (1991); Wis Stat Ann § 939.645 (West Supp 1992).

<sup>36</sup> The Supreme Court in *Mitchell* noted the precise analogy between Title VII and the hate crimes statute at issue in the case. See 113 S Ct at 2200. It is noteworthy that both

When an employer fires an employee because she is black, the government may impose sanctions without constitutional qualm. This is so even when the discharge is accomplished (as almost all discharges are) through some form of expression, for whatever expression is involved is incidental both to the act accomplished and to the government's decision to prevent it.<sup>37</sup> The analysis ought not change when a person assaults another because she is black, once again even if the conduct (assault on the basis of race) is accompanied by expression. A penalty enhancement constitutionally may follow because it is pegged to an act—a racially-based form of disadvantage—that the state wishes to prevent, and has an interest in preventing, irrespective of any expressive component. In other words, in the assault case, no less than in the discharge case, the government decides to treat race-based acts differently from similar non-race-based acts; and in the assault case, no less than in the discharge case, this decision—a decision to prevent disproportionate harms from falling on members of a racial group—bears no relation to whether the race-based act communicates a message. Thus might end the constitutional analysis.

Perhaps, however, this argument is not quite so easy as I have made it out to be. It might be said, in response, that racially-based assaults, more often than racially-based discharges, are committed in order to make a statement. If this is true, a penalty enhancement not only will restrict more speech incidentally, but also may raise a concern that the government is acting for this very purpose. Or perhaps it might be said, more generally, that the use of a discriminatory motive to define an act, even supposing the act has no expressive component, at times may be highly relevant to First Amendment analysis: consider, for example, a penalty enhance-

---

laws apply not only irrespective of whether the discrimination at issue expresses a message, but also irrespective of whether the discrimination is caused by particular beliefs. If, for example, discrimination laws prohibited discharges or assaults motivated by racial hatred—rather than simply based on race—they would pose a very different, and seemingly severe, First Amendment problem.

<sup>37</sup> Cass Sunstein makes this point in *Words, Conduct, Caste*, 60 U Chi L Rev at 827-28; his phrasing is that in such a case, the communication is merely evidence of, or a means of committing, an independently unlawful act. Professor Sunstein, however, appears to think that this analysis fails to cover hate crimes, because there the state's interest arises from the expressive nature of the conduct. As stated in the text, I do not believe this to be the case. A state has a legitimate interest in preventing, say, assaults on the basis of race, even when they are wholly devoid of expression. The interest is the same as the one in preventing discharges on the basis of race; it is an interest in eradicating racially-based forms of disadvantage generally, whether or not accompanied by communication of a message.

ment provision applicable to persons who obstruct voting on the basis of a voter's affiliation with the Republican Party.

But both of these objections seem to falter on further consideration of the nature of hate crimes regulation and the governmental interest in it. The voting obstruction law I have hypothesized (no less than a hate crimes law) applies to conduct regardless of whether it has expressive content, but the government's interest in the law always in a certain sense relates to expression: it is difficult to state, let alone give credence to, any interest the government could have, other than favoring or disfavoring points of view, for specially penalizing voting obstruction based on affiliation with a particular political party.<sup>38</sup> In the case of hate crimes laws, by contrast, the government not only is regulating acts irrespective of their expressive component, but also has a basis for doing so that is unrelated to suppressing (or preferring) particular views or expression—the interest, once again, in preventing conceded harms from falling inequitably on members of a particular racial group. In such a case, the regulation should be found to accord with First Amendment requirements, notwithstanding that it incidentally affects some expression. As the Court in *R.A.V.* noted, in referring to employment discrimination laws, “Where the government does not target conduct on the basis of its expressive content,”—and where, we might add, the government, in regulating conduct, has a credible interest that is unrelated to favoring or disfavoring certain ideas or expression—“acts are not shielded from regulation merely because they express a discriminatory idea or philosophy.”<sup>39</sup>

In accord with this reasoning, communities should be able not only to impose enhanced criminal sanctions on the perpetrators of hate crimes, but also to provide special tort-based or other civil remedies for their victims. One of the accomplishments of the anti-pornography movement has been to highlight the benefits of using the civil, as well as the criminal, laws to deter and punish undesir-

---

<sup>38</sup> The hypothetical voting law might seem very different if enhanced penalties applied to obstruction based on the voter's affiliation with any political party, rather than with the Republican Party alone. In enacting this broader law, the state could have determined that it had an interest in protecting persons from suffering disproportionate harm as a result of their political views, analogous to the interest in protecting persons from suffering disproportionate harm as a result of their race. Under the analysis suggested in the text, this new voting law would meet constitutional standards because it applies regardless whether the conduct communicates a message and because the government now has a credible interest in the law not related to favoring or disfavoring particular viewpoints and messages.

<sup>39</sup> 112 S Ct at 2546-47.

able activity.<sup>40</sup> Civil actions involve fewer procedural safeguards for the defendant, including a much reduced standard of proof; as important, they may give greater control to the victim of the unlawful conduct than a criminal prosecution ever can do. Communities therefore should consider not merely the enactment of hate crimes laws, but also the provision of some kind of "hate torts" remedies. And in determining the scope of all such laws, communities should consider the manner in which the laws apply to crimes or civil violations committed on the basis of sex, which now often fall outside the compass of hate crimes statutes.

To address the harms arising from pornography, the government has numerous available mechanisms that regulate not speech, but conduct. At an absolute minimum, states can prosecute actively, under generally applicable criminal laws, the sexual assaults and other violent acts so frequently committed against women in the making of pornography. Similarly, as Judge Easterbrook suggested in *Hudnut*, states may specifically make illegal (if they have not already) the use of fraud, trickery, or force to induce people to perform in any films, without regard to viewpoint.<sup>41</sup> Extensive regulation of such practices is the lot of many industries; the visual media surely are not entitled to any special exemption. With respect to regulatory effects of this kind too, responses based on the criminal law can be supplemented by enhanced tort remedies.<sup>42</sup>

A much more questionable means of deterring the production of pornographic works would be to press into service laws regulating prostitution, pimping, or pandering. In one recent case, an Arizona court upheld, against First Amendment challenge, the use of prostitution and pandering statutes against a woman who managed and performed in a sex show.<sup>43</sup> The court reasoned, consistent with established First Amendment doctrine, that the prosecutions were

---

<sup>40</sup> See Catharine A. MacKinnon, *Pornography, Civil Rights, and Speech*, 20 Harv CR-CL L Rev 1, 29 n 52 (1985).

<sup>41</sup> 771 F2d at 332.

<sup>42</sup> For a discussion of whether the government, in addition to banning the conduct itself, may prohibit the dissemination of speech produced by means of this unlawful conduct, see text accompanying notes 55-61.

<sup>43</sup> *Arizona v Taylor*, 167 Ariz 429, 808 P2d 314, 315-16 (1990). The state's prostitution statute prohibited "engaging in or agreeing or offering to engage in sexual conduct with another person under a fee arrangement with that person or any other person." *Id.* The use of statutes of this kind against women who merely perform in pornography raises a special concern: such prosecutions make a criminal of the very victim of exploitative practices. Moreover, these prosecutions may have little value: they are likely to deter the production of pornography far less well than prosecuting the actual pornographer under pimping, pandering, or other similar statutes, which essentially prohibit the hiring of persons to engage in sexual practices.

permissible because even if the show had expressive content, the state had acted under statutes directed at conduct in order to further interests unrelated to the suppression of expression.<sup>44</sup> The same argument could be made whenever the government acts against a pornographer under a sufficiently broad pimping or pandering statute, so long as the prosecution were based on a significant interest unrelated to speech, such as the prevention of sexual exploitation. The problem with this analysis lies in its potential scope: many films that no one would deem pornographic contain sexual conduct by hired actors and thus fall within the very same statutes. Notwithstanding all I have said above, even the neutral application of a law that is not itself about speech might in some circumstances violate the First Amendment. (Consider, to use an extreme example, an environmental law imposing a ban on cutting down trees, as applied to producers of books and newspapers.) In all probability, the use of pimping and pandering statutes in the way I have just considered suffers from this constitutional defect, given the potential for applying such statutes to large amounts of speech at the core of constitutional protection.

Those favoring the direct regulation of pornography often charge that relying exclusively on bans on conduct—most notably, a ban on coerced performances—would allow abuses currently committed in the manufacture of pornography to continue.<sup>45</sup> Such approaches, even if determinedly enforced, certainly will have less effect than banning pornography altogether. But once again, the most sweeping strategies also will be the ones most subject to constitutional challenge and the ones most subversive of free speech principles. An increased emphasis on conduct, rather than speech, provides a realistic, principled, and perhaps surprisingly effective alternative.

#### B. Viewpoint-Neutral Restrictions

The Supreme Court often has said that any speech restriction based on content, even if not based on viewpoint, presumptively violates the First Amendment.<sup>46</sup> But rhetoric in this instance is

---

<sup>44</sup> Id at 317. The key case supporting this analysis is *United States v O'Brien*, 391 US 367 (1968), in which the Court approved the use of a statute prohibiting any knowing destruction of a Registration Certificate, purportedly enacted to further the efficient operation of the draft, against a person who had burned his draft card as part of a political protest.

<sup>45</sup> See, for example, Cass R. Sunstein, *Neutrality in Constitutional Law (With Special Reference to Pornography, Abortion, and Surrogacy)*, 92 Colum L Rev 1, 23-24 (1992).

<sup>46</sup> See, for example, *Police Department of Chicago v Mosley*, 408 US 92, 95-96 (1972); *Simon & Schuster, Inc. v Members of the New York State Crime Victims Board*, 112 S Ct

semi-detached from reality. The Court, for example, sometimes has upheld regulations based on the subject matter of speech.<sup>47</sup> And the Court in several cases has approved restrictions on non-obscene but sexually explicit or scatological speech.<sup>48</sup> Cases of this kind raise the possibility of eradicating the worst of hate speech and pornography through statutes that, although based on content, on their face (and, to the extent possible, as applied) have no viewpoint bias.

One potential course is to enact legislation, or use existing legislation, prohibiting carefully defined kinds of harassment, threats, or intimidation, including but not limited to those based on race and sex. For example, in considering the St. Paul ordinance, the Court in *R.A.V.* noted that the city could have achieved "precisely the same beneficial effect" through "[a]n ordinance not limited to the favored topics"<sup>49</sup>—that is, through an ordinance prohibiting all fighting words, regardless whether based on race, sex, or other specified category. An ordinance of this kind would have presented no constitutional issue at all given the Court's prior holdings that fighting words are a form of unprotected expression.<sup>50</sup> A law prohibiting, in viewpoint-neutral terms, not merely fighting words but other kinds of harassment and intimidation would (and should) face greater constitutional difficulties, relating most notably to overbreadth and vagueness; but a carefully drafted statute might well surmount these hurdles, and such a law surely would not be subject to the selectivity analysis of *R.A.V.* Viewpoint-neutral laws of this kind—whether framed in terms of fighting words

---

501, 508-09 (1991); *Consolidated Edison Co. of New York v Public Service Commission of New York*, 447 US 530, 536 (1980).

<sup>47</sup> See, for example, *Burson v Freeman*, 112 S Ct 1846 (1992); *Greer v Spock*, 424 US 828 (1976); *CBS v Democratic National Committee*, 412 US 94 (1973). See generally Geoffrey R. Stone, *Restrictions of Speech Because of its Content: The Peculiar Case of Subject-Matter Restrictions*, 46 U Chi L Rev 81 (1978). *R.A.V.* might be thought to treat subject matter restrictions with the same distrust shown to viewpoint restrictions: the technical holding of the Court was that the St. Paul ordinance facially violated the Constitution "in that it prohibits otherwise permitted speech solely on the basis of the subjects the speech addresses." 112 S Ct at 2542. But elsewhere in the opinion, the Court made clear that its true concern related to viewpoint bias. What most bothered the Court was that the subject matter restriction operated in practice to restrict speech of only particular (racist, sexist, etc.) views. See, for example, *id.* at 2547-49.

<sup>48</sup> See *FCC v Pacifica Foundation*, 438 US 726 (1978) (indecent radio broadcast); *Young v American Mini-Theatres*, 427 US 50 (1976) ("adult" theaters); *City of Renton v Playtime Theatres, Inc.*, 475 US 41 (1986) (same).

<sup>49</sup> 112 S Ct at 2550.

<sup>50</sup> See *Chaplinsky v New Hampshire*, 315 US 568, 572 (1942). Of course, the application of the ordinance to any particular expression might well raise serious constitutional issues relating to the permissible scope of the fighting words category.

or in some other manner—might be especially appropriate in communities (such as, perhaps, educational institutions) whose very purposes require the maintenance of a modicum of decency.<sup>51</sup>

Another approach, relevant particularly to pornography, could focus on regulating materials defined in terms of sexual violence. At first glance, *R.A.V.* and (especially) *Hudnut* seem to doom such efforts, but this initial appearance may be deceptive. The problem in *Hudnut* involved the way the ordinance under review distinguished between materials presenting women as sexual equals and materials presenting women as sexual subordinates: two works, both equally graphic, would receive different treatment because of different viewpoints.<sup>52</sup> This problem, the court suggested, would not arise if a statute instead were to classify materials according to their sexual explicitness.<sup>53</sup> Indeed, the Supreme Court already has said as much by treating as non-viewpoint-based (and sometimes upholding) regulations directed at even non-obscene sexually graphic materials.<sup>54</sup> If a regulation applying to sexually explicit materials does not raise concerns of viewpoint bias, perhaps neither does a regulation applying to works that are both sexually explicit and sexually violent.

One counterargument might run that the reference to sexual violence in this hypothetical statute would function simply as a code word for a disfavored viewpoint: sexually violent materials present women as subordinates; sexually non-violent materials present women as equals; hence, the law replicates in covert language the faults of the MacKinnon-Dworkin ordinance. But this response strikes me as flawed, because many non-violent works present women as sexual subordinates, and some violent materials may not (violence is not *necessarily* a synonym for non-equality). The question is by no means free from doubt—much depends on how far the Court will or should go to find viewpoint discrimination in a facially neutral statute—but framing a statute along these lines seems worth consideration.

---

<sup>51</sup> See Robert C. Post, *Racist Speech, Democracy, and the First Amendment*, 32 *Wm & Mary L Rev* 267, 317-25 (1991), for a general discussion of the compatibility of speech regulation with the objectives of higher education.

<sup>52</sup> See 771 F2d at 328.

<sup>53</sup> *Id.* at 332-33.

<sup>54</sup> See note 48 and accompanying text. The Court has failed to indicate precisely when regulations of this kind, even assuming they are not viewpoint-based, will meet constitutional standards. All of the regulations upheld by the Court have involved not complete bans, but more limited restrictions. A law foreclosing such speech entirely would raise constitutional concerns of greater dimension.

Finally, and once again of particular relevance to pornography, the Constitution may well permit direct regulation of speech, if phrased in a viewpoint-neutral manner, when the regulation responds to a non-speech related interest in controlling conduct involved in the materials' manufacture. Assume here, as discussed above, that the government has a strong interest in regulating the violence and coercion that often occurs in the making of pornography.<sup>55</sup> Does it then follow that the government may punish the distribution of materials made in this way as well as the underlying unlawful conduct? The Supreme Court's decision in *New York v Ferber*<sup>56</sup> suggests an affirmative answer. In *Ferber*, the Court sustained a statute prohibiting the distribution of any material depicting a sexual performance by a child, primarily on the ground that the law arose from the government's interest in preventing the conduct (sexual exploitation of children) necessarily involved in making the expression. Similarly, it would appear, the government may prohibit directly the dissemination of any materials whose manufacture involved coercion of, or violence against, participants. The *Hudnut* Court specifically indicated that such a statute would meet constitutional requirements.<sup>57</sup>

Important questions remain unanswered with respect to this approach, for there are almost surely limits on the principle that the government may engage in viewpoint-neutral regulation of speech whenever it has an interest in deterring conduct involved in producing the expression. The principle itself, in addition to explaining *Ferber*, may explain such disparate outcomes as the ability of a court to enjoin the publication of stolen trade secrets and to award damages for the unapproved publication of copyrighted material.<sup>58</sup> But some hypothetical applications of the principle suggest the need for a boundary line. For example, could the government prohibit all speech whose manufacture involved violations of the Fair Labor Standards Act? Surely such a statute would violate the Constitution. Or, to use another sort of case, could the government prohibit the distribution of all national security information stolen from government agencies? An affirmative answer would require overruling the *Pentagon Papers* case.<sup>59</sup> The question arises,

---

<sup>55</sup> See text accompanying notes 41-42.

<sup>56</sup> 458 US 747 (1982).

<sup>57</sup> See 771 F2d at 332-33.

<sup>58</sup> See *Harper & Row, Publishers, Inc. v Nation Enterprises*, 471 US 539 (1985).

<sup>59</sup> See *New York Times Co. v United States*, 403 US 713 (1971). I thank Geof Stone for suggesting this example.

then, how to separate permissible from impermissible applications of the principle. I am not sure that any factor, or even set of factors, can serve to explain fully all the cases mentioned. Some relevant considerations, however, might include the value of the speech at issue, the magnitude of the harm involved in producing the speech, the extent to which prohibiting the speech is necessary to prevent the harm from occurring, and the extent to which the expression itself reinforces or deepens the initial injury.<sup>60</sup> With respect to all of these considerations, the prohibition of materials whose manufacture involves sexual violence seems similar enough to the ban in *Ferber* to suggest that the regulation, while deterring the worst forms of pornography, still would satisfy First Amendment standards.<sup>61</sup>

### C. Obscenity

The government can also regulate sexually graphic materials harmful to women by using the long-established category of obscenity. This approach to regulating such materials has come to assume the aspect of heresy in the ranks of anti-pornography feminism. Those who have argued for regulating pornography have stressed the differences, in rationale and coverage, between bans

---

<sup>60</sup> The *Ferber* Court viewed the harm involved in manufacturing child pornography as great and the value of the resulting expression as usually, though not always, slight. See 458 US at 757-58, 762-63, 773-74. With respect to the necessity of prohibiting not merely the unlawful conduct, but also the speech itself, the *Ferber* Court stated that "the distribution network for child pornography must be closed if the production of material which requires the sexual exploitation of children is to be effectively controlled." Id at 759. Finally, the *Ferber* Court noted that "the materials produced are a permanent record of the children's participation and the harm to the child is exacerbated by their circulation." Id.

<sup>61</sup> The Supreme Court's decision in *City of Renton v Playtime Theatres*, 475 US 41 (1986), might be taken to suggest—although, I believe, wrongly—a further extension of the argument: that the government may prohibit the distribution of materials even substantially correlated to unlawful conduct in manufacture, so long as the definition of these materials is viewpoint-neutral. In *Renton*, the Court upheld the regulation of adult motion picture theaters on the ground that such theaters generally correlate with a rise in crime in the surrounding neighborhood. Id at 50. The Court declined to require a showing that any particular movie theater in fact produced these results. Similarly, a statute regulating a category of speech that is highly correlated with coercion of, or violence against, women might be thought to pass constitutional muster even if a particular instance of that speech did not involve coercion or violence. This line of argument, however, takes what I believe itself to be a problematic decision much too far. Crucial to the *Renton* holding was the limited scope of the regulation under review: it zoned adult theaters, but did not prohibit them. Id at 53. A total ban on speech, based on a mere correlation between the speech and unlawful conduct (even if the conduct, as in *Renton* and here, stemmed from something other than the speech's communicative effects), would raise constitutional concerns of much greater magnitude.

on the pornographic and bans on the obscene. It is said that obscenity law focuses on morality, while pornography regulation focuses on power.<sup>62</sup> It is said that offensiveness and prurience (two of the requirements for finding a work obscene) bear no relation to sexual exploitation.<sup>63</sup> It is said that taking a work "as a whole," as obscenity law requires, and exempting works of "serious value," as obscenity law does, ill-comports with the goal of preventing harm to women.<sup>64</sup> I do not think any of this is flatly wrong, but I do wonder whether these asserted points of difference—today, even if not in the past—suggest either the necessity or the desirability of spurning the obscenity category.

My doubts began in the midst of first teaching a course on free expression. In keeping with the prevailing view, I rigidly segregated the topics of obscenity and pornography. (If I recall correctly, I taught commercial speech in between the two.) In discussing each, I iterated and reiterated the distinctions between them, in much the terms I have just described. I think I made the points clearly enough, but my students resisted; indeed, they could hardly talk about the one topic separately from the other. In discussing obscenity, they returned repeatedly to the exploitation of women; in discussing pornography, of course, they dwelt on the same. Those who favored regulation of pornography also favored regulation of obscenity—at least as a second-best alternative. Those who disapproved regulation of pornography also disapproved regulation of obscenity. Perhaps it was a dense class or I a bad teacher, but I think not; rather, I think the class understood—or, at the very least, unwittingly revealed—something important.

Even when initially formulated, the current standard for identifying obscenity was justified in part by reference to real-world harms. To be sure, the Supreme Court, in its fullest statement of the rationale for establishing the category of obscenity, spoke of the need "to protect 'the social interest in . . . morality'" and, what is perhaps the same thing, of the need "to maintain a decent society . . . ."<sup>65</sup> Here, the Court appeared to stress a version of morality divorced from tangible social consequences and related to simple sentiments of offense or disgust. But the Court also spoke

---

<sup>62</sup> See MacKinnon, *Feminism Unmodified* at 147 (cited in note 13).

<sup>63</sup> See *id.* at 174-75; Sunstein, 92 *Colum L Rev* at 20-21 (cited in note 45).

<sup>64</sup> See MacKinnon, *Feminism Unmodified* at 174-75 (cited in note 13).

<sup>65</sup> *Paris Adult Theatre I v Slaton*, 413 US 49, 59-60, 61 (1973) (emphasis deleted), quoting *Jacobellis v Ohio*, 378 US 184, 199 (1964) (Warren dissenting), and *Roth v United States*, 354 US 476, 485 (1957).

of—indeed, emphasized just as strongly—the “correlation between obscene material and crime” and, in particular, the correlation between obscene materials and “sex crimes.”<sup>66</sup> This concern too may reflect a notion of morality, but if so, it is a morality rooted in material harms.<sup>67</sup> And although some of the specific harms then perceived might now appear dated—the Court was thinking as much of unlawful acts involving “deviance” as of unlawful acts involving violence—still the Court understood the obscenity category as emerging not merely from a body of free-floating values, but from a set of tangible harms, perhaps including sexual violence.<sup>68</sup>

Much more important is the way conceptions of obscenity have evolved since then, in part because of the anti-pornography movement itself, in part because of the deeper changes that movement reflects in public attitudes and morals. This shift in understanding, I think, accounted for my classroom experience. It is hard to test a proposition of this sort, but I will hazard it anyway: one of the great (if paradoxical) achievements of the anti-pornography movement has been to alter views on obscenity—to transform obscenity into a category of speech understood as intimately related, in part if not in whole, to harms against women.<sup>69</sup> Surely, such a change in perception should come as no great surprise. It would be the more astonishing by far if obscenity were viewed today as obscenity was viewed two decades ago, when the current constitutional standard was first announced. A doctrinal test does not so easily freeze public understandings, especially when the test in part relies (as the obscenity test does) on community standards and consciousness.<sup>70</sup> Views of obscenity, in other words, are not

---

<sup>66</sup> 413 US at 58-59.

<sup>67</sup> See Daniel O. Conkle, *Harm, Morality, and Feminist Religion: Canada's New—But Not So New—Approach to Obscenity*, 10 Const Comm 105, 123-24 (1993), for discussion of these two kinds of morality (offense-based and harm-based) as reflected in obscenity doctrine.

<sup>68</sup> For this reason, I think Catharine MacKinnon's statement that obscenity is “ideational and abstract,” rather than “concrete and substantive,” represents something of an overstatement, even as applied to the initial understanding and formulation of the category. See MacKinnon, *Feminism Unmodified* at 175 (cited in note 13).

<sup>69</sup> One interesting proof (and product) of this reconceptualization is Senator Mitch McConnell's proposed legislation granting the victim of a sexual offense a right to claim damages from the distributor of any obscene work deemed to have contributed to the crime. Pornography Victims' Compensation Act of 1991, S 1521, 102d Cong, 1st Sess (Jul 22, 1991). Whatever the merits of this legislation, which raises serious concerns on numerous grounds, it clearly presupposes a link between obscenity and sexual violence.

<sup>70</sup> The obscenity standard asks whether the average person, applying contemporary community standards, would find a work prurient and offensive in its depiction of sexual

static, and they may have evolved in such a way as to link obscenity with harms to women.

Now it might be argued, in response to this claim, that so long as the formal test for determining obscenity remains the same, this reconceptualization of obscenity will avail women little, because the test's focus on prurience and offensiveness will prevent new understandings from affecting judicial outcomes. But this response seems to ignore the subtle and gradual ways law often develops. As prosecutors, juries, and judges increasingly adopt this new view of obscenity, enforcement practices and judicial verdicts naturally will come to resemble, although not to replicate, those that would obtain under an anti-pornography statute. There is in fact a substantial overlap between the categories of obscenity and pornography: most of the worst of pornography (materials with explicit and brutal sexual violence) meets the obscenity standard. As public perceptions continue to change, the application of the obscenity standard increasingly will focus on the materials causing greatest harm to women; nor need this development reflect any illegitimate acts of prosecutorial discretion.<sup>71</sup>

Moreover, this new focus may over time reshape, in a desirable manner, even the governing legal standard for determining obscenity. Doctrinal adjustments and reformulations of existing low-value categories of speech may well—and should—occur more readily than the creation of whole new categories, especially when the proposed new categories incorporate clear viewpoint bias. So, for example, the current obscenity test's requirement that materials be patently offensive may disintegrate in light of new understandings about the harms the obscenity category principally should address. This evolution of obscenity law recently has occurred in Canada, where the Supreme Court, responding to increased evidence and altered perceptions of harm to women, made sexual violence rather than sexual offensiveness the keystone of the obscenity category.<sup>72</sup> Efforts to redefine the obscenity category in this manner—a redefi-

---

conduct. It also asks whether the work lacks serious literary, artistic, political, or scientific value. See *Miller v California*, 413 US 15, 24 (1973).

<sup>71</sup> If prosecutors determine to enforce obscenity laws only against materials with a certain viewpoint, the resulting actions would be no less problematic than the MacKinnon-Dworkin statute itself. But this result is hardly the only one that could be produced by changing public norms. For example, as noted earlier and discussed again below, a focus on sexual violence arguably is not viewpoint-biased. See text accompanying notes 52-54 and 74. Thus, to the extent that prosecutors enforce obscenity laws strictly against sexually violent materials that fall within the obscenity category, their acts would not violate the R.A.V. proscription of preferring some viewpoints to others within a low-value category.

<sup>72</sup> See *Regina v Butler and McCord*, [1992] 1 SCR 452, 134 NR 81, 108-18 (Canada).

tion that, consistent with much First Amendment theory, would tend to divorce speech restrictions from simple feelings of offense—should proceed in the United States as well.<sup>73</sup>

One measure along these lines that states or localities might attempt involves the special regulation of subcategories of obscenity that contain sexual violence. *R.A.V.* might seem to bar such an approach; it held, after all, that even within low-value categories of speech, such as obscenity or fighting words, the government may not make distinctions that pose a danger of viewpoint bias. I have argued above that a statute framed in terms of sexual violence may no more implicate this principle than the several statutes upheld by the Court framed in terms of sexual explicitness.<sup>74</sup> But even if courts reject this argument, another possibility presents itself. The Court in *R.A.V.* stated as an exception to its broad rule that a subcategory of unprotected speech can be specially regulated if it presents, in especially acute form, the concerns justifying the exclusion of the whole category from First Amendment protection.<sup>75</sup> It is hard to know what this exception means, especially in light of the Court's refusal to apply it to the category of race-based fighting words, which appears to pose in especially acute form the dangers giving rise to the entire fighting words category. It is no less difficult to determine what the exception *should* mean, given the ability to characterize in many different (and even conflicting) ways the concerns underlying any low-value category and the ease of restating those concerns with respect to any given subcategory. But given the Court's acknowledgment of the relationship between sexual crimes and obscenity, some consideration should be given to whether a statute focusing on the particular kinds of obscenity that most contribute to sexual violence would or should fall within the *R.A.V.* exception.<sup>76</sup>

---

<sup>73</sup> It might be argued that such a redefinition of the obscenity category would render it viewpoint-based and therefore inconsistent with the First Amendment. This argument depends first on the proposition that a statute framed in terms of sexual violence is viewpoint-based, which I have discussed in the text accompanying notes 52-54. As important, the argument depends on the proposition that the obscenity category is not now viewpoint-based—in other words, that it does not now constitute some kind of exception to the rule of viewpoint neutrality. This proposition is difficult to maintain given the obscenity test's reliance on community standards of offensiveness. See Sunstein, 92 Colum L Rev at 28-29 (cited in note 45). As between an obscenity doctrine that focuses on sexual prurience and offensiveness and an obscenity doctrine that focuses on sexual prurience and violence, the former would appear to pose the greater danger of viewpoint bias.

<sup>74</sup> See text accompanying notes 52-54 and notes 71 and 73.

<sup>75</sup> 112 S Ct at 2545-46.

<sup>76</sup> The Court wrote, for example, that "a State may choose to regulate price advertising in one industry but not in others, because the risk of fraud (one of the characteristics of

The key point here is that regulation of obscenity may accomplish some, although not all, of the goals of the anti-pornography movement; and partly because of the long-established nature of the category, such regulation may give rise to fewer concerns of compromising First Amendment principles. Even for those who think that the obscenity doctrine is in some sense a second-best alternative, it represents the first-best hope of achieving certain objectives. And the obscenity doctrine itself may benefit by transformative efforts, as these efforts bring the doctrine into greater accord with the harm-based morality of today, rather than of twenty years ago.

#### D. Exceptions to Viewpoint Neutrality

The final approach I will discuss, although far more briefly than it deserves, involves crafting arguments to support explicit exceptions to the rule against viewpoint discrimination for pornography or hate speech. As noted earlier, exceptions to this rule do exist, but without any clear rationale; the Court, in upholding viewpoint discriminatory actions, simply has ignored their discriminatory nature. We know, from the decision in *R.A.V.* and the affirmation of *Hudnut*, that the Court will follow no such course of studied inattention with respect to pornography or hate speech: in both cases, the presence of viewpoint discrimination was considered—and was declared dispositive. The question, then, arises: Is it possible to make a convincing argument to the contrary? Is it possible, that is, to accept viewpoint neutrality as a general principle, but to support an exception to that principle either for pornography or for hate speech? The challenge here is to explain in credible fashion what makes one or two or three viewpoints (or one or two or three instances of viewpoint discrimination) different from all others—sufficiently different to support an exception and sufficiently different to ensure that the exception retains “exceptional” status. I cannot here provide the answer to that question. Instead, I will confine myself to some general observations about what considerations might be relevant to the inquiry.

Two factors necessary (but, I will argue, generally insufficient) for departing from the norm of viewpoint neutrality are (1) the

---

commercial speech that justifies depriving it of full First Amendment protection) . . . is in its view greater there.” *Id.* at 2546. So too, it might be said, a State may choose to regulate in a special manner sexually violent obscenity because it poses a greater risk of contributing to sexual crimes—one of the characteristics of obscenity that justifies depriving it of full First Amendment protection.

seriousness of the harm the speech causes, and (2) the "fit" between the harm and the viewpoint discriminatory mechanism chosen to address it. The first consideration has an obvious basis: to the extent a viewpoint causes insignificant harm, the state's decision to suppress that viewpoint must rest not on legitimate reasons but on mere dislike of the idea at issue. The second consideration is related and not much more mysterious: when the government restricts a viewpoint, but the viewpoint is not coextensive with the harm allegedly justifying the governmental action, we may wonder (once again) whether the action is in fact motivated by simple distaste for the message. I have no doubt that a regulation of pornography and hate speech would satisfy the first inquiry, and little doubt that such a regulation could be carefully enough constructed to satisfy the second. Is that, however, sufficient?

I think not. Assume, for example, a carefully crafted regulation of abortion advocacy, counseling, or referral (the category of speech involved in *Rust v Sullivan*<sup>77</sup>), designed to reduce the incidence of abortions. Proponents of the regulation might urge that the law is precisely crafted to reduce the significant harms stemming from abortion; hence the law satisfies the two inquiries set forth above. I presume this outcome would strike many as irretrievably wrong. But, some opponents of the regulation might contend, the example fails to prove my larger point because the "harms" in the hypothetical case (however serious some might find them) are in fact widely contested and for that reason cannot form the basis of viewpoint regulation. These opponents might contrast a precisely crafted regulation of pro-smoking speech, designed to reduce the frequency of tobacco use. In that case, the harms are not contested; hence the regulation can go forward. The contrast here has much intuitive appeal, and I am not at all sure it has nothing to teach us. But this general line of reasoning makes the protections of the First Amendment weakest at the very point where views are the most unorthodox and unconventional. And even if I am wrong to think this result upside-down and unacceptable, another question would follow: Are not the harms caused by pornography and hate speech—characterized most generally as racial and sexual subordination—also very much contested? If they were not, the debate over hate speech and pornography might not have reached so intense a level.

---

<sup>77</sup> 111 S Ct 1759, 1765 (1991).

Assuming, then, that harm and "fit" cannot alone justify viewpoint discrimination, perhaps the addition of low-value speech can do so. In other words, if legislators can make the case that speech leads to harm, if the speech regulated correlates precisely with that harm, and if the speech is itself low-value, then any viewpoint discrimination involved in the regulation becomes irrelevant.<sup>78</sup> At first glance, of course, *R.A.V.* definitively rejected this argument: the very holding of that case was that even within a low-value category of speech, viewpoint discrimination is generally prohibited. So, to use one of the Court's hypotheticals, the government may proscribe libel, but may not proscribe only libel attacking the government; or, to use something near the actual case, the government may prohibit fighting words, but may not prohibit only racist fighting words.<sup>79</sup> But what, then, are we to make of a category like obscenity—an entire low-value category (rather than a subdivision thereof) that seems to incorporate some viewpoint bias?<sup>80</sup> Could it possibly be the case that viewpoint discrimination built into the very definition of a low-value category is permissible, whereas viewpoint discrimination carving up a neutrally defined low-value category is not?

The proposition is perhaps less silly than it appears, for the latter, but not the former, lacks the precise "fit" that I above termed necessary for viewpoint regulation. When the Court establishes a low-value category, such as obscenity, it determines that the harms caused by the covered speech so outweigh its (minuscule) value that regulation of the speech, even if viewpoint discriminatory, will be permitted. The Court, in effect, predecides that regulation of the entire category will arise not from governmental hostility to the ideas restricted, but rather from a neutral decision based on harms and value; the viewpoint bias will occur as a mere byproduct of the fact that only the restricted ideas cause great harms and have sparse value. This predetermination insulates the government from a charge of viewpoint bias when the government regulates the entire category. But the establishment of a low-value category has no such effect when the government regulates within the category on the basis of a viewpoint extraneous to the cate-

---

<sup>78</sup> I take Cass Sunstein to be making something like this argument in these pages. See 60 U Chi L Rev at 829 (cited in note 17).

<sup>79</sup> 112 S Ct at 2543 & n 4. The actual ordinance, as construed, prohibited race-based fighting words (discriminating by subject matter), but the Court argued that this restriction operated in practice in the same way as an ordinance banning racist fighting words (discriminating by viewpoint). See *id.* at 2547-48.

<sup>80</sup> See notes 13 and 73.

gory's boundaries. In that case, there is reason to suspect that the government is acting not for the reasons already found by the Court to be legitimate, but rather out of hostility to a message. The critical failure in such a regulation relates to "fit": because the regulation is underinclusive—because it does not regulate all speech previously determined to cause great harm and have no value—the concern arises that the government has an illegitimate motive. Hence, to say, as the Court did in *R.A.V.*, that the government may not engage in unrelated viewpoint discrimination within a low-value category—may not, for example, ban only obscenity produced by Democrats—is not to say that viewpoint may not enter into the very definition of a low-value category. Once again, in the latter case viewpoint serves as a placeholder for a balance of harms and values found legitimate by the Court; in the former case, viewpoint serves as a warning signal that the government is acting for other reasons.

But even if this distinction holds, the hard question remains: should the Court accept pornography or hate speech as a low-value category of expression? The currently recognized categories of low-value speech seem to share the trait, as Cass Sunstein writes, that they are neither "intended [nor] received as a contribution to social deliberation about some issue."<sup>81</sup> That definition offers several lessons for any regulation, concededly based on viewpoint, either of hate speech or of pornography. In the case of hate speech, such an ordinance should be limited to racist epithets and other harassment: speech that may not count as "speech" because it does not contribute to deliberation and discussion. In the case of pornography, any ordinance should be limited to materials that operate primarily (as obscene materials operate primarily) as masturbatory devices; in addition, an explicit exception, like that in the obscenity standard, for works of serious value ought to be incorporated. Only if pornography and hate speech are defined in this narrow manner might (or should) the Court accept them as low-value categories—a classification that, it must be remembered, depends at least as much on the non-expressive quality of the speech as on the degree of harm the speech causes.

In addition to all this, perhaps one other factor—the modesty, or limited nature, of the viewpoint restriction—should be considered prior to recognizing a low-value category of speech incorporating viewpoint bias. This inquiry would focus on whether the regu-

---

<sup>81</sup> Sunstein, 60 U Chi L Rev at 807.

lation of the category wholly excises the viewpoint from the realm of public discourse or cuts off only a limited means of expressing the viewpoint.<sup>22</sup> Even the MacKinnon-Dworkin version of anti-pornography legislation would do only the latter: it would prohibit not all messages of sexual subordination, but only those messages expressed in a sexually graphic manner. This feature seems critical to the establishment of any exception to the viewpoint neutrality principle. The broader the restriction, the more it will skew public discourse toward some views and away from others. And the larger the skewing effect, the greater the chances of improper governmental motivation; a wholesale, more than a marginal, restraint suggests a government acting not for neutral reasons, but out of simple hostility to the idea restricted. Of course, the inquiry into the scope of a viewpoint restriction does not lend itself to scientific precision. The matter is always one of degree, involving the drawing of a line someplace on a spectrum. The inquiry, too, is complicated by the issue whether the particular means restricted (even if technically modest) constitute the most effective way of delivering the message, such that the restriction ought to be treated as sweeping. But the haziness of the endeavor does not gainsay the need to engage in it. For a viewpoint restriction that results in excising ideas from public discourse ordinarily ought not to be countenanced—even when the restriction applies only to low-value speech and even when the restriction closely responds to serious harms.

#### CONCLUSION

The presumption against viewpoint discrimination, relied upon in *Hudnut* and further strengthened in *R.A.V.*, has come to serve as the very keystone of First Amendment jurisprudence. This presumption, in my view, has real worth, in protecting against improperly motivated governmental action and against distorting effects on public discourse. And even if I assign it too great a value, the principle still will have to be taken into account by those who

---

<sup>22</sup> I do not at all advocate here that courts consider the modesty of a viewpoint restriction in all cases involving viewpoint regulation. Rather, I mean that courts should ask this question when the other criteria, discussed above, for departing from the viewpoint neutrality rule have been met. This approach is similar to the one used in *City of Renton v Playtime Theatres, Inc.*, 475 US 41, 53 (1986), in which the Supreme Court looked to the scope of the speech restriction at issue—an inquiry the Court normally eschews—in a case involving low-value speech. For a detailed discussion generally disapproving any inquiry into the modesty of a viewpoint restriction, although not considering the precise issue raised here, see Stone, *Content Regulation and the First Amendment* at 200-33 (cited in note 26).

favor any regulation either of hate speech or of pornography. I have suggested in this Essay that the regulatory efforts that will achieve the most, given settled law, will be the efforts that may appear, at first glance, to promise the least. They will be directed at conduct, rather than speech. They will be efforts using view-point-neutral classifications. They will be efforts taking advantage of the long-established unprotected category of obscenity. Such efforts will not eradicate all pornography or all hate speech from our society, but they can achieve much worth achieving. They, and other new solutions, ought to be debated and tested in a continuing and multi-faceted effort to enhance the rights of minorities and women, while also respecting core principles of the First Amendment.

THE CHANGING FACES OF FIRST  
AMENDMENT NEUTRALITY: *R.A.V. v*  
*ST. PAUL, RUST v SULLIVAN*, AND THE  
PROBLEM OF CONTENT-BASED  
UNDERINCLUSION

Consider two cases—the most debated, as well as the most important, First Amendment cases decided by the Supreme Court in the past two Terms: *R.A.V. v St. Paul*,<sup>1</sup> invalidating a so-called hate speech ordinance, and *Rust v Sullivan*,<sup>2</sup> upholding the so-called abortion gag rule. On their face, the cases have little in common; certainly, the Justices deciding them saw no connection. Yet just underneath the surface, the cases have a similar structure, implicate an identical question, and fall within a single (though generally unrecognized) category of First Amendment cases. Along with many other cases to which neither has been assimilated, *R.A.V.* and *Rust* are, on this level, essentially the same—except that the one issue of First Amendment law they posed was answered by the Court in two different ways.

The equation of the cases at first glance is jarring, because an

---

Elena Kagan is Assistant Professor of Law, The University of Chicago Law School.

AUTHOR'S NOTE: I am grateful to Akhil Amar, Anne-Marie Burley, Richard Epstein, Abner Greene, Larry Lessig, Michael McConnell, Richard Posner, Jed Rubenfeld, and Cass Sunstein for very helpful comments on earlier drafts of this article. The Bernard G. Sang Faculty Fund and the Bernard Meltzer Fund provided financial assistance for this project.

<sup>1</sup> 112 S Ct 2538 (1992).

<sup>2</sup> 111 S Ct 1759 (1991).

orthodox understanding of First Amendment law highlights only the cases' dissimilarities. On such a view, the Court in *Rust* faced the new—and exceedingly difficult—First Amendment problem of selective funding of speech by the government.<sup>3</sup> The question was whether the federal government could fund a range of family planning services, but exclude from such funding abortion counseling, advocacy, or referral. Call this a selective subsidization question or call it an unconstitutional conditions question,<sup>4</sup> the essential nature of the inquiry is the same: it focuses on the government's ability to influence the realm of speech by distributing its own (wholly optional) largesse. By contrast, according to the orthodox view, the Court in *R.A.V.* faced the classic—and largely settled—First Amendment problem of the outright prohibition of a certain kind of speech by the government. The question was whether a municipality could criminalize the use of "fighting words" that provoke violence "on the basis of race, color, creed, religion, or gender." The focus was on the ability of the government to ban speech on the basis of content through use of the government's coercive power. Seen in this light, *Rust* and *R.A.V.* raised different problems, and it is no wonder that the cases provoked divergent responses: a stark rejection of the First Amendment claim in *Rust*, a powerful affirmation of the First Amendment claim in *R.A.V.*<sup>5</sup>

<sup>3</sup> To call such questions "new" is in a significant sense to compress history. The potential for these questions to emerge has existed in great measure since the rise of the regulatory state, and the Court has decided a number of First Amendment cases involving selective subsidization issues during the past decades. See, for example, *Speiser v Randall*, 357 US 513 (1958). Indeed, even prior to the creation of the regulatory state, issues of this kind could arise in such contexts as government property or employment. See, for example, *McAuliffe v Mayor of New Bedford*, 155 Mass 216, 29 NE 517 (1892). That these issues are still considered in any degree novel may have as much to do with their intractability—with the continuing inability of courts and commentators to resolve them—as with their timing.

<sup>4</sup> Phrased in the language of conditions, the question is whether the government could condition its grant of funding on the content of the recipient's speech.

<sup>5</sup> The variance—and, I will soon argue, the inconsistency—in the Court's responses to *Rust* and *R.A.V.* goes yet further than that suggested in the text. Four of the five Justices who voted to deny the First Amendment claim in *Rust* voted to sustain a broad First Amendment position in *R.A.V.* Those four were Chief Justice Rehnquist and Justices Scalia, Kennedy, and Souter; of the *Rust* majority, only Justice White rejected the broad First Amendment argument in *R.A.V.*, though concurring in the result on narrower grounds. Conversely, the two active Justices who wished to sustain the First Amendment claim in *Rust* (Justices Blackmun and Stevens) rejected the *R.A.V.* majority's broad First Amendment reasoning, though again concurring in the result. Justice O'Connor, who voted with the concurring Justices in *R.A.V.*, declined to take a position on the constitutional question in *Rust*, and in the interim between the two cases Justice Thomas, who joined the *R.A.V.* majority, replaced Justice Marshall, who joined the *Rust* dissent.

But is this the only—is this the best—way to view these cases? Or can they be recast—the issues in them redescribed—so that an underlying similarity leaps out? A few preliminaries at once suggest themselves. First, both cases involve speech of a particularly controversial—many believe deeply harmful—kind. That abortion advocacy is the bane of a certain segment of the political right and that racist speech is the bane of a certain segment of the political left must be considered, for First Amendment purposes, not a distinction, but a core likeness. Next, in each case the government responded to this controversy by engaging in a form of content discrimination, disfavoring certain substantive messages as compared to others. Both cases thus raise general questions of First Amendment neutrality: whether, when, and how the government may tip the scales for (or against) certain messages—or, stated otherwise, to what extent the government is required, with respect to the content of speech, to play a neutral role. But more than this must be said to assimilate the cases, for surely the question of First Amendment neutrality may present itself in different contexts, and different contexts may demand different approaches and legal rules. The key, then, to understanding the connection between *R.A.V.* and *Rust* is to note that in both cases, the issue of neutrality arises in the same way—that in both, the structure of the problem is the same.

How is this so? Briefly stated for now, *Rust* and *R.A.V.* both raise the question: If, in a certain setting, the government need not protect or promote any speech at all, may the government choose to protect or promote only speech with a certain content? *Rust* is easily seen in this light. The government, we believe, is not constitutionally required to promote speech through the use of federal funds.<sup>6</sup> May the government then fund whatever speech it wants? Or does it face constraints in selectively promoting expression? The question is similar in *R.A.V.* The government is not constitutionally required to tolerate any “fighting words” at all. May the government then permit some but not all fighting words? Or is it constitutionally constrained from selectively doling out this favor? The question posed in each case is in an important sense the question of First Amendment neutrality in its starkest form:

---

<sup>6</sup> There are exceptions to this widely accepted principle. See note 53. Yet the rule remains generally valid and served as the foundation for *Rust*.

when speech, considered broadly, has no claim to government promotion or protection, what limitations does government face in voluntarily advancing some messages, but not all?

This issue, which I will call the issue of content-based underinclusion, extends far beyond *Rust* and *R.A.V.* themselves. It links a wide variety of First Amendment cases and defines a largely unacknowledged First Amendment category. The question arises in cases involving selective funding of speech (such as *Rust*), selective prohibition of wholly proscribable speech (such as *R.A.V.*), selective bans on speech in non-public forums, and selective imposition of otherwise valid time, place, or manner restrictions (which may or may not involve the use of government property). At present, some of these cases—most notably, those involving funding decisions—are viewed as raising nasty, even intractable issues; others are seen as far more transparent. But if we recognize that all belong to one broad category, we may come to doubt our certainty as to some, even as we may gain guidance on others.

In this article, I view *R.A.V.* and *Rust* as reflecting on each other and, together, as reflecting on a broader range of First Amendment cases. My purpose is to elucidate connections that the Court's discourse has obscured, to explore what turns out to be a far-flung problem, and to essay some steps toward a solution. In Part I, I summarize the opinions in *R.A.V.* and *Rust*, showing how the majority opinion in *R.A.V.* echoes the principal dissent in *Rust* and how the majority opinion in *Rust* anticipates the principal concurrence in *R.A.V.* In Part II, I provide a fuller statement of the structural congruity of the cases and the issue they present, and I connect them with other kinds of First Amendment cases raising the question of content-based underinclusion. Part III considers two objections to this broad linkage: one based on the distinction between penalties and nonsubsidies, the other based on what appears to be the plenary power of the government to engage in speech itself. Finally, Part IV offers some tentative thoughts on the resolution of the problem of content-based underinclusion.

## I

*R.A.V.* arose from the City of St. Paul's decision to charge a juvenile under the St. Paul Bias-Motivated Crime Ordinance for allegedly burning a cross on the property of an African-American

family. The ordinance, as written, declared it a misdemeanor for any person to “place[] on public or private property a symbol, object, appellation, characterization or graffiti, including, but not limited to, a burning cross or Nazi swastika, which one knows or has reasonable grounds to know arouses anger, alarm or resentment in others on the basis of race, color, creed, religion or gender . . . .”<sup>7</sup>

The trial court dismissed the charge on the ground that the St. Paul ordinance was overbroad. The Minnesota Supreme Court reversed, holding that the ordinance, as properly construed, banned only expression not protected by the First Amendment. The court relied on *Chaplinsky v New Hampshire*, which declared that “fighting words”—defined as words “which by their very utterance inflict injury or tend to incite an immediate breach of the peace”—could be punished without “rais[ing] any constitutional problem.”<sup>8</sup> According to the Minnesota Supreme Court, the St. Paul ordinance was constitutional because it extended only to expression that fell within the *Chaplinsky* formulation (although, of course, not to all such expression): the law covered “fighting words” that injured or provoked violence on the basis of race, color, creed, religion, or gender.<sup>9</sup>

All nine Justices agreed to strike down the ordinance as construed by the Minnesota Supreme Court, but none pretended to have achieved anything more than surface unanimity. Four of the Justices invalidated the law only because, in their view, the Minnesota Supreme Court had failed in its attempt to limit the ordinance to expression proscribable under *Chaplinsky*; the ordinance thus remained overbroad.<sup>10</sup> The majority declined to consider this argument, and the real controversy in the case lay elsewhere. It centered on the following question: Assuming the St. Paul ordinance

---

<sup>7</sup> Minn Stat § 292.02 (1990).

<sup>8</sup> 315 US 568, 572 (1942).

<sup>9</sup> See *In re Welfare of R.A.V.*, 464 NW2d 507, 510–11 (1991).

<sup>10</sup> In holding that the St. Paul ordinance reached only “fighting words” as defined by *Chaplinsky*, the Minnesota Supreme Court had suggested that the *Chaplinsky* definition included expression that by its very utterance caused (in the words of the St. Paul ordinance) “anger, alarm or resentment.” 112 S Ct at 2559. The four concurring Justices objected to this sweeping understanding of *Chaplinsky*. The Justices stated, in accord with other post-*Chaplinsky* decisions, that the fighting words doctrine articulated in that case in no way allowed the restriction of speech that inflicted only such “injury” as “hurt feelings, offense, or resentment.” *Id.*

reached only expression proscribable under *Chaplinsky*, did the ordinance remain invalid because it reached some, but not all, of this expression—because it banned, on the basis of content, only certain fighting words?

Justice Scalia, writing for the majority,<sup>11</sup> answered the question in the affirmative and invalidated the ordinance on this ground. In prior cases, Justice Scalia readily admitted, the Court had made a judgment that fighting words could be banned entirely—a judgment based on the view that such words are “‘of such slight social value as a step to truth that any benefit that may be derived from them is clearly outweighed by the social interest in order and morality.’”<sup>12</sup> The Court even had gone so far as to say that fighting words and other similar categories of expression are “‘not within the area of constitutionally protected speech’” and that the “‘protection of the First Amendment does not extend’” to them.<sup>13</sup> But were these statements to be taken as “literally true”?<sup>14</sup> Did the First Amendment vanish from the landscape because the government had no obligation to permit the utterance of fighting words? Not at all.

What remained fixed on the constitutional terrain was an obligation of content-neutrality, perhaps slightly relaxed in the context of proscribable speech, but still with significant bite. No matter, for example, that the government may proscribe libel; “it may not make the further content discrimination of proscribing *only* libel critical of the government.”<sup>15</sup> No matter that a city may ban obscenity; it may not “prohibit . . . only that obscenity which in-

<sup>11</sup> The majority also included Chief Justice Rehnquist and Justices Kennedy, Souter, and Thomas.

<sup>12</sup> 112 S Ct at 2543 (quoting *Chaplinsky*, 315 US at 572). Justice Scalia’s opinion nowhere questioned the fighting words doctrine as formulated in *Chaplinsky*; that doctrine was treated throughout the opinion as a given. It is conceivable that some unstated discomfort with the fighting words doctrine contributed to, or even caused, the *R.A.V.* decision; on this view, the reasoning of the Court in *R.A.V.* operated as a kind of second-best surrogate for the ideal but seemingly intemperate course of overruling the doctrine entirely. Cf. Richard A. Epstein, *Foreword: Unconstitutional Conditions, State Power, and the Limits of Consent*, 102 Harv L Rev 4, 28–31 (1988) (explaining various prohibitions on selective government action found in unconstitutional conditions cases as a second-best means of constraining unwisely granted government power). I assume here that the *R.A.V.* Court meant what it said and that its rationale was something more than a pretext for limiting a doctrine it did not like, but felt bound to tolerate.

<sup>13</sup> Id at 2543 (quoting, inter alia, *Roth v United States*, 354 US 476, 483 (1957), and *Base Corp. v Consumers Union*, 466 US 485, 504 (1984)).

<sup>14</sup> Id.

<sup>15</sup> Id (emphasis in original).

cludes offensive political messages.”<sup>16</sup> Similarly, with respect to the case at hand: no matter that a city may bar all fighting words; it may not (as, the majority held, St. Paul did) bar only those fighting words addressing a particular subject or expressing a particular viewpoint.<sup>17</sup> Although the category of fighting words is “unprotected”—although it has, “in and of itself, [no] claim upon the First Amendment”—the government does not have free rein to regulate selectively within the category.<sup>18</sup> Even wholly proscribable categories of speech are not “entirely invisible to the Constitution, so that they may be made the vehicles for content discrimination.”<sup>19</sup> To sustain all content discrimination within categories of speech, simply because the categories as a whole are proscribable, would be to engage in “a simplistic, all-or-nothing-at-all approach to First Amendment protection . . . at odds with common sense.”<sup>20</sup>

Justice White, in a concurring opinion,<sup>21</sup> took direct issue with this reasoning: for him, the only relevant fact was that fighting words as a category could be banned under the First Amendment. Once the determination had been made that fighting words generally had no claim to First Amendment protection, the conclusion followed that the government could regulate such expression freely—even if that regulation took the form of content discrimination. “It is inconsistent to hold that the government may proscribe an entire category of speech . . . but that the government may not treat a subset of that category differently without violating the First Amendment; the content of the subset is by definition . . . undeserving of constitutional protection.”<sup>22</sup> Indeed, such a holding foolishly would force the government to choose between regulating all proscribable speech or none at all.<sup>23</sup> In Justice White’s frame-

---

<sup>16</sup> *Id.* at 2546 (emphasis deleted).

<sup>17</sup> *Id.* at 2547.

<sup>18</sup> *Id.* at 2545.

<sup>19</sup> *Id.* at 2543.

<sup>20</sup> *Id.*

<sup>21</sup> Justice White’s opinion was joined in full by Justice Blackmun and Justice O’Connor. Justice Stevens joined only the portion of the opinion stating that the ordinance was overbroad; he specifically rejected both Justice White’s and Justice Scalia’s approaches to the question discussed in the text. I discuss aspects of Justice Stevens’s opinion in Part IV.

<sup>22</sup> *Id.* at 2553.

<sup>23</sup> In this manner, Justice White was able to throw back upon Justice Scalia the charge of all-or-nothingism. See *id.* Justice Stevens charged both opinions with manifesting that apparently discredited approach to First Amendment questions. See *id.* at 2562, 2567.

work, when speech had no claim to constitutional protection, government selectivity made no First Amendment difference;<sup>24</sup> if the government had no obligation to permit fighting words at all, then it faced no constraints in permitting some fighting words but not others.

Turn now to *Rust*, and compare the structure of the argument. The Department of Health and Human Services had issued regulations governing the allocation and use of Title X grants.<sup>25</sup> These regulations prohibited Title X-funded projects from providing abortion counseling or referrals (instead requiring them to provide referrals for prenatal care), as well as from encouraging, promoting, or advocating abortion. Title X grantees challenged the regulations, alleging (among other claims) that they violated the First Amendment.<sup>26</sup> The grantees argued in part that, by virtue of the regulations, the availability of subsidies now hinged on the content of speech—or, more specifically, its viewpoint: the government would subsidize a wide range of speech on family planning and other topics (including anti-abortion speech), but not abortion counseling, referral, or advocacy.

A majority of the Court, speaking through Chief Justice Rehnquist, rejected this argument. The starting point, for the Court, was that the Constitution required no subsidization of speech at all: “[A] legislature’s decision not to subsidize the exercise of a fundamental right does not infringe the right.”<sup>27</sup> For the majority it followed that the government could also subsidize speech selec-

---

<sup>24</sup> Justice White stated that the Equal Protection Clause, as distinct from the First Amendment, would pose a barrier to differential treatment not rationally related to a legitimate government interest. See *id.* at 2555. Ahkil Amar suggests that in acknowledging the relevance of the Equal Protection Clause, Justice White may have conceded the crucial point: that even within the realm of unprotected speech, some state action is illegitimate. See Ahkil R. Amar, *Comment: The Case of the Missing Amendments: R.A.V. v. City of St. Paul*, 106 Harv L Rev 124, 130 & n 46. The question remains, though: Exactly what state action is illegitimate? Justice White’s rational basis test, which would strike down legislation “based on senseless distinctions,” 112 S Ct at 2556 n 9, will not lead to the same results as Justice Scalia’s demanding First Amendment scrutiny.

<sup>25</sup> Such grants are made under Title X of the Public Health Service Act, 42 USC §§ 300–300a-6 (1988), which provides monies for family planning services. The HHS regulations appear at 42 CFR §§ 59.7–59.10 (1991).

<sup>26</sup> The grantees also argued that the regulations failed to comport with the governing statute and that they violated the Fifth Amendment right of women to choose to have an abortion. The Court rejected both these claims.

<sup>27</sup> 111 S Ct at 1772 (quoting *Regan v Taxation with Representation*, 461 US 540, 549 (1983)).

tively within broad limits:<sup>28</sup> the Court had rejected the proposition “that if the government chooses to subsidize one protected right, it must subsidize analogous counterpart rights.”<sup>29</sup> In effect, the “general rule that the Government may choose not to subsidize speech” implied a corollary: the government may choose which speech to fund.<sup>30</sup> And what of the usual First Amendment proscription against viewpoint discrimination? The Chief Justice suggested that in this context the term had no application: when the government “has merely chosen to fund one [speech] activity to the exclusion of the other[,]” the government “has not discriminated on the basis of viewpoint.”<sup>31</sup> In allotting funds, the government was entitled to make “value judgment[s].”<sup>32</sup> The government could subsidize speech promoting democracy, but not speech promoting fascism;<sup>33</sup> the government could subsidize speech of family planning clinics (including anti-abortion speech) except for abortion advocacy and referral. All followed from a simple point: “Title X subsidies are just that, subsidies.”<sup>34</sup> The statement echoes Justice White in *R.A.V.*: Fighting words are just that, fighting words. When the government has no general obligation, it has no obligation of neutrality.

Justice Blackmun’s dissent in *Rust* vigorously disputed this proposition. Justice Blackmun acknowledged that the government generally has a choice whether to fund the exercise of a constitutional right, but he insisted that “there are some bases upon which gov-

---

<sup>28</sup> Noting that funding by the government might not “invariably [be] sufficient to justify government control over the content of expression,” the Court proposed two potential exceptions: when the subsidy was offered to a university or when the subsidy took the form of providing a public forum. *Id.* at 1776.

<sup>29</sup> *Id.* at 1773.

<sup>30</sup> *Id.* at 1776.

<sup>31</sup> It is conceivable that the Chief Justice intended to make a far narrower point than that suggested in the text: he may have meant only that the particular funding decision at issue did not involve viewpoint discrimination (as generally understood in First Amendment law), because the HHS regulations merely drew a distinction, on the basis of subject matter, between speech concerning pre-conception family planning and all other speech. In one portion of the opinion, the Court indeed approaches this argument. See *id.* at 1772. But the argument, aside from being fallacious in light of the language of the regulations, see text at note 99, cannot be thought to represent the whole, or even a major part, of the Court’s reasoning: so narrow an interpretation of the decision makes most of the *Rust* opinion, including the statements emphasized in the text, incomprehensible.

<sup>32</sup> *Id.* at 1772.

<sup>33</sup> *Id.* at 1773.

<sup>34</sup> *Id.* at 1775 n. 5.

ernment may not rest [a] decision” to fund expression.<sup>35</sup> Selective funding becomes impermissible when based upon the content—most clearly, upon the viewpoint—of the expression. The government may not “discriminate invidiously in its subsidies” of speech by basing them on ideological viewpoint.<sup>36</sup> Thus, Justice Blackmun concluded, “[t]he majority’s reliance on the fact that the Regulations pertain solely to funding decisions simply begs the question.”<sup>37</sup> The point echoes Justice Scalia in *R.A.V.*: The concurrence’s reliance on the fact that the St. Paul ordinance pertains solely to fighting words simply begs the question. Even in this circumstance, the government retains an obligation of neutrality.

Thus do the arguments in *Rust* and *R.A.V.* mirror each other. Between the two cases, the Court switched sides: the dissent in *Rust* became the *R.A.V.* majority, the majority in *Rust* became a concurrence in *R.A.V.* So too did most of the individual Justices trade positions; the difference in the outcome of the cases is hardly due to the change of mind of a single Justice.<sup>38</sup> But the structure of the dispute in the two cases is almost precisely the same. And that is because the *Rust* Court and the *R.A.V.* Court faced the same issue—a distinctive kind of First Amendment neutrality issue, extending far beyond *R.A.V.* and *Rust* themselves, which might best be labeled content-based underinclusion.

## II

What, precisely, is content-based underinclusion? Suppose that the government, consistent with the First Amendment, may limit—by prohibiting or by refusing to subsidize—either an entire category of speech or all speech within a particular context. Now suppose that the government declines to go so far: rather than limiting speech to the full extent of its constitutional power, the government chooses to limit only some expression—and that on the basis of content. The resulting government action is, in the ordinary sense, narrower than the action stipulated to be constitutional. That is, the merely partial limitation allows more expres-

---

<sup>35</sup> *Id.* at 1781.

<sup>36</sup> *Id.* at 1780 (quoting *Regan*, 461 US at 548); see *id.* at 1782.

<sup>37</sup> *Id.* at 1781.

<sup>38</sup> See note 5.

sion. Yet this “narrower” action incorporates a content-based distinction: it picks and chooses among expression on the basis of what is said. The question thus becomes whether and when a government that has the power to restrict speech generally may instead limit select kinds of expression. Or, looked at from a different angle, the question is whether the government may voluntarily promote or protect some (but not all) speech on the basis of content, when none of the speech, considered in and of itself, has a constitutional claim to promotion or protection.

Such underinclusion—government may ban all speech in a category, but instead bans only some, defined by content—is a particular kind of content-based restriction, by no means equivalent to all government actions falling within the broad content-based category.<sup>39</sup> In many—indeed, most—cases of content-based speech restrictions, the question of inequality between different kinds of expression is wrapped in, and in practice inseparable from, a theoretically distinct issue: the permissibility of the burden placed on the speech affected. Consider, for example, a case arising from a statute that criminalizes in all contexts constitutionally protected speech—say, seditious advocacy. In deciding such a case, the Court usually will not ask whether the government has a sufficient reason to treat speech of one kind (seditious advocacy) differently from speech of another; rather, the Court will ask merely whether the government has a sufficient reason to restrict the speech actually affected.<sup>40</sup> The framing of the inquiry relates to the nature of the problem: in such a case, the issue is not underinclusion, for the government could not cure the constitutional flaw by extending the restriction to all speech regardless of its content.

By contrast, in a content-based underinclusion case, equality is

---

<sup>39</sup> Justice Scalia attempts in *R.A.V.* to avoid the term “underinclusiveness” in favor of the broader term “content discrimination,” apparently because he thinks the former term more liable to the concurring opinions’ charges of First Amendment absolutism. See 112 S Ct at 2545. But content-based underinclusion is no more than a distinctive kind of content-based distinction, and analysis explicitly focusing on underinclusion (when it exists) does no more than respond to the peculiar nature of the governmental action and the peculiar concerns it raises. Justice Scalia himself recognizes the need to distinguish among different kinds of content-based distinctions when he concedes that content-based analysis may take a somewhat different form in the context of wholly proscribable speech than in other First Amendment contexts. See *id.*

<sup>40</sup> See, e.g., *Brandenburg v Ohio*, 395 US 444 (1969) (per curiam); see generally Geoffrey R. Stone, *Content Regulation and the First Amendment*, 25 Wm & Mary L Rev 189, 202–3 (1983).

all that is at issue. Here, the Court usually will state the issue in terms of (and only in terms of) equal treatment. The Court will ask not whether the government has a sufficient reason for restricting the speech affected (taken in isolation), but whether the government has a sufficient reason for restricting the speech affected *and* not restricting other expression.<sup>41</sup> Once again, the framing of the inquiry follows from the structure of the problem. In these cases, by definition, the restriction is permissible but for the inequality, and the constitutional infirmity thus may be erased by extending the restriction to additional speech as well as by eliminating it entirely.<sup>42</sup> The First Amendment functions in these cases solely as a guarantee of some kind of equality on the plane of content.

The issue of content-based underinclusion arises in many settings—all superficially unlike, but all essentially similar.<sup>43</sup> One set of cases presenting the issue involves the selective imposition of otherwise reasonable time, place, or manner restrictions. Assume that a city may ban the use of noisy soundtrucks between sunset and sunrise in residential districts. Now assume that the city, rather than enacting this flat ban, exempts the use of soundtrucks to laud city government. One approach to this law holds that the burden imposed on speech is itself constitutionally permissible, but strikes down the law because of the content-based exemption.<sup>44</sup>

---

<sup>41</sup> On occasion the Court has focused on differential treatment without stating that a generally applied restriction of the same kind would be constitutional. But in almost all of the cases in which the Court has framed the question in this manner, such a general restriction on speech at least arguably would have satisfied constitutional standards. See, for example, *Police Dep't v Mosley*, 408 US 92 (1972).

<sup>42</sup> Justices frequently object to the Court's analysis in such cases precisely on the ground that it permits the enactment of a broader speech restriction. See 112 S Ct at 2553 (White concurring); *id* at 2561–62 (Stevens concurring); *Metromedia, Inc. v San Diego*, 453 US 490, 564 (1981) (Burger dissenting); *Carey v Brown*, 447 US 455, 475 (1980) (Rehnquist dissenting).

<sup>43</sup> See Geoffrey R. Stone, Louis M. Seidman, Cass R. Sunstein, and Mark V. Tushnet, *Constitutional Law* at 1337–62 (Little, Brown, 2d ed 1991), which organizes some cases along the lines I suggest in a section entitled “Equality and Free Expression.”

<sup>44</sup> A court also might take either of two different approaches to the law. First, a court might ask whether the government has a compelling reason to burden the speech affected, without any exploration of the scope of the exemption. Under this approach, the content-based exemption serves to heighten the standard of review (to one of compelling interest); the ultimate inquiry, however, remains focused on the permissibility of the burden imposed, irrespective of the exemption. Second, a court might again focus on the permissibility of the burden imposed, but use the exemption not merely to heighten the standard of review, but to discredit the justification for the general speech restriction. For example, in the hypothetical given, a court might reason that if the city allows this exemption, then the city

Under this analysis, the permissibility of the general restriction is irrelevant: the government, even when it has discretion over allowing speech at all, may not grace a certain kind of speech with its special favor.<sup>45</sup>

Many Supreme Court cases reviewing limited time, place, or manner regulations incorporate this understanding of the content-based underinclusion problem and the analysis associated with it. In some of these cases, the regulations applied to the use of public forums. For example, in *Police Dept. v Mosley*,<sup>46</sup> the Court reviewed an ordinance that prohibited picketing on public streets near a school during certain hours, but exempted labor picketing from the general restriction. The Court held the ordinance unconstitutional because of the distinction between labor picketing and other picketing—because the ordinance worked a content-based “selective exclusion from a public place.”<sup>47</sup> In other cases, the time, place, or manner restriction has applied outside the realm of public property. Thus, in *Metromedia v San Diego*,<sup>48</sup> the Court considered the

---

must view the interest in quiet during evening hours as insignificant, in which case the general restriction must fall. An analysis of this kind, although relying heavily on the exemption, in the end tests the constitutionality of the actual burden imposed on speech and finds that burden excessive. In other words, the exemption itself is not what is invalid; rather, the exemption proves the invalidity of a more general ban. See Stone, *Content Regulation and the First Amendment*, 25 Wm & Mary L. Rev at 202–7 (cited in note 40).

<sup>45</sup> The Court in *R.A.V.* itself recognized the link between *R.A.V.* and cases of the kind discussed in the text. The Court compared the proscription of fighting words to the proscription of a noisy soundtruck. See 112 S Ct at 2544–45. The analogy implies that content-based distinctions within a generally proscribable category of speech (such as fighting words) present the same question as content-based distinctions superimposed on an otherwise valid time, place, or manner regulation.

<sup>46</sup> 408 US 92 (1972).

<sup>47</sup> Id at 94; see also *Carey v Brown*, 447 US 455 (1980) (invalidating on the same ground a statute prohibiting all picketing, except labor picketing, on streets surrounding residential places). *City of Lakewood v Plain Dealer Publishing Co.*, 108 S Ct 2138 (1988), presented the same issue in a different form. The case involved standards governing the allocation of city permits for newspaper vending machines. All assumed that the provision of city property (even public forum property) for vending machines was wholly optional, in the sense that the city government could choose whether it wished to allow any machines at all. The majority held that if the city chose to exercise this power, it must do so under standards that would safeguard against content discrimination. The dissent, written by Justice White and closely resembling his opinion in *R.A.V.*, concluded that because the First Amendment did not obligate the city to allow the placement of newsracks on city streets (or, in his words, because the placement of newsracks—like the use of fighting words—was not “protected by the First Amendment”), the city had no obligation to promulgate protective standards. In *Lakewood*, however, even Justice White agreed that were the city actually to engage in content discrimination in allocating newsrack permits, the First Amendment would come into play.

<sup>48</sup> 453 US 490 (1981).

legality of an ordinance restricting the use of billboards unless they fell within certain categories defined by content, such as political campaign signs or signs indicating the temperature or time. Here too, the Court struck down the law on the basis of its selectivity, entirely independent of the extent of the burden that the law imposed on the covered speech. The message in these cases, regardless whether public property was involved, was the same: even if speech generally may be regulated through reasonable time, place, or manner restrictions, such restrictions may not be imposed on speech only of a certain content.

All of these cases thus concern the same issue as *Rust* and *R.A.V.*, although they reach results identical only to the latter. In *Rust*, the Court permitted the government to favor (through funding) certain kinds of speech, on the ground that the government need not have favored any. In *Mosley* and *Metromedia*, the Court refused to allow the government to engage in similar selectivity: to favor (through donating public property or granting a regulatory exemption) certain kinds of speech on the ground that all speech could have been disfavored. If anything, as I will later discuss, *Rust* might be thought to raise a graver First Amendment problem, because the selectivity there was based on viewpoint, whereas in *Mosley* and *Metromedia*, it was based (at least facially) only on subject matter. In any event, the cases raised the same essential issue: the demands of First Amendment neutrality in a sphere in which government action respecting speech is in the first instance optional.

The Court often confronts the identical issue—but handles it differently—when dealing with speech restrictions applicable to non-public forums. Within broad limits, the government may choose to impose in such places sweeping restrictions on speech, so long as generally applicable.<sup>49</sup> Depending on the nature of the non-public forum, the government may have discretion to ban speech entirely. Frequently, however, the government chooses to restrict—in this context, up to the point of banning altogether—

---

<sup>49</sup> Restrictions must be “reasonable” in light of the nature and purposes of the non-public forum, but this standard frequently allows even wholesale prohibition of speech. For an example of the ease with which the reasonableness standard may be met in the context of non-public forums, see *International Society for Krishna Consciousness, Inc. v Lee*, 112 S Ct 2701 (1992). By contrast, in a public forum (whether traditional or designated), the government has only very narrow discretion to curtail speech generally, through limited time, place, or manner restrictions.

only speech of a certain content. Thus, the question once more arises: in circumstances in which the government need not allow or foster any speech, may it decide to allow or foster some speech on the basis of content?

Two cases will serve to illustrate how the issue arises—and how the Court has handled it—in this context. In *Lehman v City of Shaker Heights*,<sup>50</sup> the Court reviewed a municipal policy of refusing to sell advertising space on city buses to persons who wished to use the space to engage in political speech. After finding that the advertising space did not constitute a public forum, and thus that no general right of access applied, the Court was left with the question whether the municipality could bar only a certain kind of speech. Similarly, in *Greer v Spock*,<sup>51</sup> the Court considered whether a military base, also a non-public forum, could bar speeches and demonstrations of a partisan political nature, while allowing other kinds of expression. In these cases and others,<sup>52</sup> the Court has permitted some content-based distinctions (including those based on subject matter), but has drawn the line at distinctions that are based on viewpoint. The government may not use its broad discretion over the property it owns to advantage some viewpoints at the expense of others, but as in *Lehman* and *Greer* may make other distinctions based on content.

These cases too resemble *Rust* and *R.A.V.*, except in the rules the Court has established and the results it has reached. Banning all fighting words, as in *R.A.V.*, is no more problematic than banning all speech in a non-public forum. Yet in *R.A.V.*, the Court invalidated selective proscription, suggesting that even subject-matter distinctions violated the First Amendment, whereas in *Lehman* and *Greer*, the Court upheld such selective proscription. Perhaps, as I shall later discuss, the cases may be distinguished by virtue of the kind of content discrimination in each. But surely it should make no difference that the one case involves a selective ban within a wholly proscribable category of speech, the others a selective ban within a non-public forum. In both, what is at issue is the ability of the government to restrict some (but not all) speech

<sup>50</sup> 418 US 298 (1974).

<sup>51</sup> 424 US 828 (1976).

<sup>52</sup> See *Perry v Perry*, 460 US 37 (1983) (upholding statute granting preferential access to an interschool mail system); *Cornelius v NAACP*, 473 US 788 (1985) (upholding government policy limiting access to a charity drive aimed at federal employees).

when the government has the discretion to restrict the speech entirely.

From the discussion so far, it may come as little surprise to discover that even within a single setting—that of selective funding decisions—the problem of content-based underinclusion has bedeviled the Court. The government, as a general rule, need not fund any speech, whether through direct expenditures, tax exemptions, or other mechanisms.<sup>53</sup> But what if the government chooses to fund some (but not all) speech on the basis of content? Prior to *Rust*, the Court had confronted on several occasions this issue of selectivity in public funding decisions. In *Arkansas Writers' Project v Ragland*, for example, the Court considered the constitutionality of extending a tax exemption to religious, professional, trade, and sports journals, but not to general-interest magazines.<sup>54</sup> The Court struck down the exemption scheme because it rested on content distinctions, even though turning only on subject matter. In *Regan v Taxation with Representation*, by contrast, the Court approved a congressional decision to grant a tax subsidy to veterans' organizations, but not to other organizations, engaged in lobbying efforts.<sup>55</sup> There, the Court indicated (as it has in the non-public forum cases) that only viewpoint-based selectivity in government funding would violate the First Amendment.<sup>56</sup> Finally, as discussed

---

<sup>53</sup> This general rule is burdened with at least one prominent exception. The government has a broad obligation to permit speech in public forums; this donation of property for speech purposes is a form of funding. In addition, the government may have a duty to provide police protection and like services to speakers in certain circumstances. See *Edwards v South Carolina*, 372 US 229, 231–33 (1963); *Cox v Louisiana*, 379 US 536, 550 (1965). Once again, in providing these services, the government effectively funds expression. See generally Owen M. Fiss, *Why the State?*, 100 Harv L Rev 781, 786 (1987); Cass R. Sunstein, *Free Speech Now*, 59 U Chi L Rev 255, 273–74 (1992).

<sup>54</sup> 481 US 221 (1987).

<sup>55</sup> 461 US 540 (1983).

<sup>56</sup> The debate in *Ragland* and *Regan*, as in most such cases, focused explicitly on the question whether the government's power to refuse all funding implied a power to fund selectively. In dissent in *Ragland*, Justice Scalia saw as dispositive "the general rule that 'a legislature's decision not to subsidize the exercise of a fundamental right does not infringe the right.'" 481 US at 236 (quoting *Regan*, 461 US at 549). In *Regan*, the majority expounded this reasoning, citing the discretion of Congress over "this sort of largesse" and the absence of any First Amendment right to subsidization of speech. 461 US at 549. Other cases presenting substantially the same issue, in the context of government provision of services, are *Board of Education v Pico*, 457 US 853 (1982), in which the Court disapproved the removal of specified books from a school library over the objection that the government had no constitutional obligation to make available any book in a library, and *Southeastern Promotions v Conrad*, 420 US 546 (1975), in which the Court disapproved the exclusion of the musical "Hair" from a city auditorium over the objection that the city had substantial discretion to determine the nature of the entertainment it wished to support.

previously, the Court in *Rust* suggested that in the funding context even the prohibition on viewpoint discrimination does not apply: the discretionary nature of funding decisions obviates any requirement of government neutrality among different kinds of expression.

What appears to emerge from the cases I have discussed—*Rust*, *R.A.V.*, and all the rest—is a set of diverse and contradictory responses to a single (and ubiquitous) First Amendment problem. All these cases, I have argued, pose the issue of content-based underinclusion, and yet the Court has failed to recognize this essential sameness. The argument, however, is so far only half complete. For although I have stated what bonds the cases, I have not yet explored what might be thought to unglue them. Perhaps there are real differences among these cases—distinctions that reseparate in a principled manner what I have grouped together.

### III

In this Part, I consider two objections to the proposition that *Rust* and *R.A.V.* belong to a single category of cases in which the government engages in content-based underinclusion. The first objection turns on the distinction between penalties and nonsubsidies, familiar from the Court's treatment of unconstitutional conditions cases. Cases such as *Rust*, it is said, involve nonsubsidies, whereas cases such as *R.A.V.* involve penalties; and selectivity with respect to nonsubsidies, but not penalties, is permissible. But the distinction between nonsubsidies and penalties founders in cases involving content-based underinclusion; perhaps more important, even if the distinction could be drawn, it would have no significance within this set of cases.

The second objection to viewing these cases as part of a single category relies on the government's plenary power to engage in speech itself. If the government has power to speak unrestrictedly, the argument runs, so too does the government have uncurtailed power to hire "agents" to engage in speech activities: thus does the government action in a case like *Rust*, but not in a case like *R.A.V.*, receive constitutional approval. But this approach also overlooks the distinctive character of content-based underinclusion cases, here by misunderstanding the way in which government action in these cases relates to the government's own expression. Both approaches fail to distinguish *Rust* and *R.A.V.*; both fail to fracture

the category of content-based underinclusion; both fail to answer the question of First Amendment neutrality that category poses.

A

At the base of *Rust* lies the view that nonsubsidies and penalties are different—different in the sense that they can be distinguished from each other, and different also in the sense that the distinction matters. The government may not “penalize” a person for engaging in abortion advocacy, but the government may refuse to “subsidize” such speech, even if it subsidizes other, competing expression. The distinction between nonsubsidies and penalties runs across the gamut of unconstitutional conditions cases, whether or not involving the First Amendment; in these cases, the most common approach is to label governmental actions as either a penalty or a nonsubsidy, to declare the former coercive and unconstitutional, to declare the latter noncoercive and constitutionally permitted.<sup>57</sup>

This distinction prompts an obvious response to the argument I have been making. In discussing *Rust*, *R.A.V.*, and other cases, I have formulated the issue at stake in something like the following way: When may the government permit or subsidize some (but not all) speech on the basis of content in circumstances in which it need not permit or subsidize any? A skeptic might claim that the disjunctives in this statement are doing all the work—in other words, that I am conflating, through these simple “or”s, two separate inquiries. One question (raised, for example, by *Rust*) involves selective subsidies; the other (raised, for example, by *R.A.V.*) involves selective penalties. In that distinction, the argument further runs, lies a critical difference.

A first response to this argument contests the ease—or even the coherence—of an effort to sort out penalties from nonsubsidies in any content-based underinclusion case. In funding cases such as *Rust*, government action that seems to be a mere nonsubsidy becomes a penalty if viewed from a different, and no more contestable, perspective. Less obviously, the same is true (in reverse) of non-funding cases involving underinclusion, such as *R.A.V.*: gov-

---

<sup>57</sup> See, for example, *Regan*, 461 US 540; *Harris v McRae*, 448 US 297 (1980); *Speiser v Randall*, 357 US 513 (1958).

ernment action that seems, intuitively, a penalty becomes a mere nonsubsidy with a similar change in perspective.

Consider first a selective funding case like *Rust*, in which the difficulty of drawing the penalty/nonsubsidy distinction has frequently been noted.<sup>58</sup> In refusing to provide grants for abortion referrals, is the government penalizing or merely declining to subsidize this exercise of First Amendment rights? The answer rests upon the choice of a position—to use the inevitable jargon, a baseline—from which to measure the action. If the starting point assumes an absence of funding for any family planning services, including abortion referral, then the government action at issue is a nonsubsidy. If, by contrast, the starting point assumes funding for all family planning services, including abortion referral, then the government decision is a penalty.

The difficulty in such cases arises from the task of determining which position to adopt given that the action occurs within a realm of (frequently exercised) government prerogative. Presumably, the government action at issue should be viewed from the position of whatever state of affairs—funding or non-funding—is in some sense normal or natural. But in a world in which the government may and frequently does fund private speech and other activity, but has no general constitutional obligation to do so, the choice of this position is by no means obvious. What is the normal or natural state of affairs in such a world? Stated otherwise, what is a citizen (here, a family planning provider) entitled to expect? Nothing? Something? If the latter, what? The answers frequently are elusive.

Perhaps less obviously, the same difficulties attend any attempt to categorize the governmental action at issue (as penalty or non-subsidy) in a case like *R.A.V.* A direct prohibition of speech, backed by sanctions, might seem the archetypal penalty. But the question in an underinclusion case, such as *R.A.V.*, is in fact more complicated. Remember that the government, acting within the Constitution, either may permit or may ban fighting words; the First Amendment has nothing to say respecting that decision. If that is so, we may measure the government action at issue from either of two perspectives. We may assume a perspective in which

---

<sup>58</sup> See, for example, Seth F. Kreimer, *Allocational Sanctions: The Problem of Negative Rights in a Positive State*, 132 U Pa L Rev 1293 (1984); Kathleen M. Sullivan, *Unconstitutional Conditions*, 102 Harv L Rev 1413 (1989); Cass R. Sunstein, *Why the Unconstitutional Conditions Doctrine Is an Anachronism*, 70 BU L Rev 593 (1990).

the government tolerates all fighting words; in that case, the prohibition of racial fighting words indeed smacks of a penalty. But alternatively, we may assume a perspective in which the government prohibits all fighting words; in that case, a ban on racial fighting words seems a mere nonsubsidy (with any exemption from the general prohibition counting as a subsidy).

As in the funding cases, the choice between the two stances—protection of fighting words or no protection of fighting words—is frequently unclear, and for much the same reason. Given a world in which the government may (and frequently does) but need not protect fighting words, either stance may seem justified. In this context too, it is no mean feat to determine the normal or natural state of affairs, or a citizen's entitlement. And thus in this context too, it is no mean feat to characterize the government action at issue as either a penalty or a nonsubsidy.

Consider, for example, two alternative avenues that a municipality might take to achieve the result of the St. Paul ordinance. First, suppose that a city government initially outlawed all fighting words and then, at some later date, repealed the measure except as to racial fighting words. The repealer in this example is as optional as the provision of funds in *Rust*. It follows that the remaining prohibition, no less than the refusal to fund abortion advocacy, can be considered a mere nonsubsidy. Or, second, suppose that a city government enacted a statute prohibiting fighting words generally, but then exempting, as a special act of legislative grace, non-racial fighting words. Here too, an obvious argument can be made that the exemption is a subsidy, all else nothing more than a refusal to subsidize.

This characterization seems more natural in the hypothetical cases than in *R.A.V.* itself, but that in no way undermines the point I am making. The characterization seems more apt because in choosing a stance from which to view government action, we instinctively consider how the world looked prior to the action and whether the action singles out certain speech for favorable or unfavorable treatment.<sup>59</sup> But this is—or, at the very least, should

---

<sup>59</sup> See Kreimer, 132 U Pa L Rev at 1359–71 (cited in note 58). Kreimer explicitly advocates the use of these factors to classify government action as a penalty or a nonsubsidy and to determine, on the basis of this classification, the action's constitutionality. My own proposed analysis does not depend on these considerations because it views as essentially irrelevant in the underinclusion context the determination whether government action constitutes a penalty or subsidy. See text following note 64.

be—as true in funding cases as in non-funding underinclusion cases such as *R.A.V.* What the hypothetical cases show is that the same debate over the proper characterization of government action may arise in each of these contexts.

Thus far, the discussion suggests two points: first, that cases like *R.A.V.* and *Rust* cannot easily be distinguished on the ground that the one involves a penalty, the other a subsidy; and second, that the distinction fails because, as shown previously, the cases alike emerge from an area of government discretion. Lest it be at all unclear, I emphasize that I am not, either here or elsewhere in this essay, equating funding cases with all cases involving a direct prohibition of speech. Rather, I am equating funding cases with a specific kind of non-funding case—that involving underinclusion. In these cases, as in funding cases, classification of the government action at issue (as penalty or nonsubsidy) is problematic. It is so because these cases, like funding cases, arise against a backdrop of government prerogative: government may, but need not, act with respect to the speech at issue. Were the Constitution to command a certain action, the problem would evaporate. If the First Amendment, say, required the government to protect fighting words, the requirement itself would establish the proper baseline, and any deviation from the protection of fighting words would constitute a penalty. Similarly in the funding cases, if the Constitution required the government to pay for the exercise of speech rights, any refusal to fund speech would penalize the speaker. The difficulty arises when government has no such general obligation—when (assuming no breach of applicable neutrality requirements) it can protect or not protect, fund or not fund as it chooses.

The essential point applies well beyond the particular contexts of *Rust* and *R.A.V.* As we have seen, general government prerogative exists in a number of First Amendment contexts: not only when the government decides whether to fund speech (*Rust*), or to ban speech falling within proscribable categories (*R.A.V.*), but also when the government decides whether to prohibit speech in non-public forums, as in *Greer*, or to issue reasonable time, place, or manner regulations, as in *Mosley*. Here too we may ask whether the government, in allowing only non-political speech on an army base, has penalized political speech or subsidized non-political speech. Or whether the government, in permitting only labor speech around a school during certain hours, has granted a subsidy to labor speech or imposed a penalty on all other expression.

In all of these underinclusion cases, we may play out endless arguments about whether government action with respect to some (but not all) speech has subsidized or penalized; we may say that the government has subsidized expressive activities in declining to exercise the full powers allotted to it under the First Amendment, or we may say that the government has penalized expressive activities in exercising only some subset of those powers. What alone is clear is that the subsidy/penalty line, properly understood, fails to separate any one of the contexts involving content-based underinclusion from the others. If one can be classified as a mere subsidy case, so too can they all.

The argument so far, however, seems subject to the objection that it disregards the ordinary meaning of the terms "subsidy" and "penalty." In common parlance, to subsidize speech means to pay for it; the government subsidizes expression when it picks up the costs of such activity, transferring them from a speaker to taxpayers generally. By contrast, to penalize speech means to impose a burden on a speaker—by fine or other means—that extends beyond requiring her to pay for her own expression.<sup>60</sup> From this standpoint, *Rust* involves a subsidy because the government is paying for speech (thus redistributing from taxpayers to speaker), whereas *R.A.V.* involves a penalty because the government is imposing an extra cost on the speaker (thus effectively redistributing in the opposite direction). Therein, it might be said, lies the difference.<sup>61</sup>

A bit of examination, however, reveals otherwise. The reason is simple: There are many ways for the government to pay for speech,

---

<sup>60</sup> Richard Epstein and Michael McConnell, in slightly different ways, build their conceptions of the whole unconstitutional conditions doctrine on this redistributive conception of the subsidy/penalty distinction (although McConnell also believes that some government actions counting as subsidies under this analysis still may violate the First Amendment). See Epstein (cited in note 12); Michael W. McConnell, *Unconstitutional Conditions: Unrecognized Implications for the Establishment Clause*, 26 San Diego L Rev 255 (1989).

<sup>61</sup> Under this approach, some "funding" cases of course will turn out to involve penalties, rather than subsidies. One example is *FCC v League of Women Voters*, 468 US 364 (1984), in which the Court invalidated a statute prohibiting broadcasters who received any federal monies from airing editorials; the effect of the statute was not merely to cut off government funding of editorials (a nonsubsidy under this approach), but to cut off funding of all the broadcaster's activities if it aired editorials (a penalty under this approach because the benefits withheld went beyond the costs of the speech). See Michael W. McConnell, *The Selective Funding Problem: Abortions and Religious Schools*, 104 Harv L Rev 989, 1016–17 (1991). The primary point I will make is different: that "non-funding" underinclusion cases like *R.A.V.* may turn out to involve subsidies under a test focusing on whether government is merely refusing to pay for speech or exacting some additional cost from the speaker.

and all content-based underinclusion cases—regardless whether they involve the writing of a check from tax revenues—involve some mechanism by which the government picks up some of the costs of a speaker's expression.

Consider in this regard the ordinance in *R.A.V.*, which regulated a brand of fighting words. Such expression, by definition, imposes a cost not merely on other individuals (the targets of the fighting words), but on society at large: fighting words “are of such slight social value as a step to truth that any benefit that may be derived from them is clearly outweighed by the social interest in order and morality.”<sup>62</sup> It is indeed partly because of the social cost caused by fighting words that the Court has placed them in a wholly proscribable category. May it then not be said that in declining to regulate fighting words, the government picks up the cost of the speech, effectively paying (or forcing other citizens to pay) for it? The regulation of fighting words then appears a mere nonsubsidy, the refusal to regulate a classic example of subsidization.<sup>63</sup> Under this approach to the penalty/subsidy distinction, there is no more a constitutional “penalty” on speech in *R.A.V.* than there was in *Rust*. Both involve decisions to subsidize some expressive activities and not others.

Other kinds of content-based underinclusion cases also raise, in this sense, the issue of selective subsidization. Return here to the non-public forum cases such as *Greer*, which involved speech on a military base. The donation of such public property—property whose ordinary use is to some extent incompatible with expression—constitutes a subsidy, an absorption by the public of the costs associated with allowing expressive activity in the forum. The

---

<sup>62</sup> *Chaplinsky v New Hampshire*, 315 US 568, 572 (1942). The cost of fighting words may take a number of forms. If such words “by their very utterance inflict injury,” they will at least impose a direct harm on their target; if they “tend to incite an immediate breach of the peace,” they will impose as well a cost on the general public, including money spent for police protection.

<sup>63</sup> The same is true of the regulation of speech falling within any other category of wholly or partially proscribable expression, such as obscenity or some kinds of libel. Such regulation appears a mere nonsubsidy, in that it operates to prevent the speaker from transferring significant costs to the public; conversely, a refusal to regulate in these areas works as a subsidy, with the public determining to absorb the costs of the expression. For discussions of the way in which constitutional privileges in libel law subsidize speakers at the expense of those defamed, see Richard A. Posner, *Economic Analysis of Law* § 27.2 at 670 (Little, Brown, 4th ed 1992); Frederick Schauer, *Uncoupling Free Speech*, 92 Colum L Rev 1321, 1326–43 (1992).

denial of access to such property, by contrast, appears as a simple refusal to subsidize expression.<sup>64</sup> The same is true of cases arising from selective imposition of otherwise valid time, place, or manner restrictions, such as *Metromedia*. Here too, the government has determined that speech (in the form of billboards) imposes costs on the public. With respect to certain kinds of speech, that cost is absorbed by the taxpayers; with respect to other kinds of speech, the cost is thrown back on the speaker.

The ability to view all underinclusion cases in this manner again springs from their common grounding in a sphere of government discretion. As a general rule, the government has discretion to regulate or limit speech (assuming no violation of neutrality principles) precisely when such regulation plausibly may be described as a mere nonsubsidy in the sense just described. Thus, even if we view the subsidy/penalty line as appropriately defined by the direction of redistribution (from the speaker to the public or from the public to the speaker), cases such as *R.A.V.*—cases in which the government starts with general discretionary powers—appear not very different from direct funding cases like *Rust*. Whatever differences may exist in the form of the subsidy cannot be thought of constitutional significance.

But more than this may be said, for even if the penalty/subsidy distinction could serve to separate some underinclusion cases from others (*Rust*, for example, from *R.A.V.*), the distinction would remain, in the context of underinclusion cases, essentially irrelevant. Assume for the moment that the action involved in *R.A.V.* constitutes a “penalty.” The First Amendment objection to the action cannot focus on the penalty itself—cannot focus, for example, on the extent to which it, relative to a subsidy, cuts off speech—given that the fighting words doctrine permits the government to penalize

---

<sup>64</sup> The relation of this analysis to public forum doctrine raises interesting questions. As previously noted, the government has a broad obligation to donate public forums for expressive purposes. The public forum cases thus might be viewed as stating an exception to the general rule that the government need not subsidize expression; indeed, I have considered public forums as forced subsidies at note 53. In keeping with the understanding of subsidies and penalties used in this discussion, however, we might consider the public forum cases not to involve subsidies at all. If public forums are at least in part defined as places compatible with expressive activity, then permitting speech in such places imposes few additional costs on the public. Cf. McConnell, *The Selective Funding Problem*, 104 Harv L Rev at 1033 (cited in note 61). This case, however, becomes more difficult to make as public forums are increasingly defined, as they have been in recent years, simply in terms of some historical criteria. See *International Society for Krishna Consciousness v Lee*, 112 S Ct 2701 (1992).

all speech of this kind. The objection instead must turn on government selectivity: the government has (dis)avored some speech on illegitimate grounds. In other words, if a selective penalty in a case like *R.A.V.* is constitutionally forbidden, the reason must have everything to do with the selection, and nothing to do with the penalty, which is, in and of itself, perfectly permissible. And if this is so, any distinction between a case like *R.A.V.* and a case like *Rust* cannot lie in the differing terms “penalty” and “subsidy.” These terms should be viewed as constitutionally irrelevant; what has meaning in the cases—and in all underinclusion cases—is government selection. The Court’s focus should be on this issue, and not on a set of terms bearing no real relation to it. The penalty/subsidy distinction provides meager aid in explaining *Rust*, *R.A.V.*, or any other case of content-based underinclusion.

## B

Unstated in any decision, but perhaps vaguely perceived by the Justices, is another notion—this one relating to the government’s own speech—that may explain the divergent outcomes in *Rust* and *R.A.V.* and, more broadly, challenge the existence of a single category of content-based underinclusion cases encompassing *Rust*, *R.A.V.*, and others. The argument starts from the premise—not undisputed but generally accepted—that the First Amendment places few limits on the government’s own expressive activities; by and large, the government may speak as it chooses.<sup>65</sup> Of course, as a physical if not a constitutional matter, “the government” cannot speak; it can speak only through employees and agents. To say, then, that the First Amendment allows the government to speak is to say that the First Amendment allows the government (more precisely, its employees and agents) to hire employees and agents to do its speaking for it.<sup>66</sup>

---

<sup>65</sup> For purposes of this discussion, I accept the premise that the First Amendment imposes only minor limits on the government’s own speech. For a lengthy and critical exploration of this premise, see Mark G. Yudof, *When Government Speaks* (University of California Press, 1983).

<sup>66</sup> The Supreme Court has indicated that the First Amendment protects even an individual’s decision to hire or otherwise pay for a speaker, but also has suggested that the constitutional interest in such vicarious speech is of some lesser magnitude than the interest in direct speech. See *Buckley v Valeo*, 424 US 1 (1976) (discussing why a limitation on contributions to political campaigns poses fewer constitutional problems than a limitation on direct campaign expenditures).

From this premise emerges a claim that (at least some) government funding cases differ from all other cases of content-based underinclusion. When the government funds speech, even of hitherto private parties, the government is merely hiring agents to engage in speech for it. In paying for speech, it is speaking; if the latter is permissible, so is the former. Thus a decision like *Rust* becomes justifiable: in funding certain kinds of speech, the government effectively is engaging in the speech, and so the Constitution imposes few limits. But the same cannot be said, or so the argument goes, of a case like *R.A.V.*, which involves restrictions on the speech of private parties. The government's plenary power over its own speech provides a constitutional basis for decisions to fund expression of a particular kind, but provides no basis for decisions, even if wholly voluntary, to permit speech of a certain content.<sup>67</sup>

This argument can be contested on two independent grounds. The first disputes the equation of "government speech" and government funded speech. The second disputes the differentiation, with respect to "government speech," of funding decisions and other kinds of content-based underinclusion.

To appreciate some of the difficulties involved in equating government speech with government funding—because government can speak, it can fund others to speak—consider the following hypothetical: a city council enacts an ordinance providing that any person who endorses the actions of city government shall be entitled to a cash grant or tax exemption.<sup>68</sup> The city government itself—by which I mean municipal employees acting in their official capacity—constitutionally could engage in speech of this kind, and such speech might drown out, and hence render ineffective, countervailing expression. Given this power to speak, the hypothetical subsidy scheme cannot be attacked on the bare ground that it skews public debate about municipal government; the government's own speech also may have a skewing effect. And yet, the hypothetical

---

<sup>67</sup> I am grateful to my colleague Michael McConnell for raising this argument with me, though I do not think it should be taken (at least in this barebones form) as a statement of his position.

<sup>68</sup> Few would question the equivalence of a cash grant or other direct expenditure and a tax exemption, deduction, or credit in a scheme of this kind. As the Supreme Court has recognized, "Both tax exemptions and tax deductibility are a form of subsidy that is administered through the tax system." *Regan v Taxation with Representation*, 461 US 540, 544 (1983). Indeed, such tax provisions frequently are referred to as "tax expenditures." See Bernard Wolfman, *Tax Expenditures: From Idea to Ideology*, 99 Harv L Rev 491 (1985).

funding scheme seems (at the least) constitutionally problematic—far more so than what might be called direct government speech. The First Amendment problems also seem severe in a case, more closely analogous to *Rust*, in which the government makes cash grants not to the public at large, but to all political clubs for purposes of speech endorsing city government. Why do these funding programs appear to present greater constitutional difficulties than the government's own expression?<sup>69</sup>

As an initial matter, when the government itself speaks in favor of a position, we (the people) know who is talking and can evaluate the speech accordingly. (When the government speaks to laud itself, we may pay the speech little attention.) By contrast, when the government finances hitherto private parties to do its speaking, we may have little understanding of the source of the expression. This problem is particularly acute if we do not know of the existence of the funding scheme; then we will consistently mistake the interested for the impartial. But even if we know of the funding scheme, we will face a problem of attribution. The speakers may have engaged in the same expression without any government funding; alternatively, the speakers may have foregone their expression (or even espoused a different view) in the absence of a subsidy. We do not know whether to treat the speakers as independent or as hired guns. We thus may give the speech more (or less) weight than it deserves.

A related concern is that the funding scheme will operate to distort or influence the realm of private expression in a manner that systemically advantages public power. When the government speaks directly, it merely adds a voice (though perhaps a resounding one) to a conversation occurring among private parties. When the government speaks through subsidy schemes, it may change and reshape the underlying dialogue. What once were private choices—shall I praise the city government, criticize it, or say nothing at all?—now become in some measure governmental, as citizens calculate a set of economic incentives offered to them by government actors. The resulting choices by private individuals and organizations may give greater volume to the government's voice than the government could have achieved on its own. As

---

<sup>69</sup> For a related discussion of this question, see Cass R. Sunstein, *The Speech Market* (Free Press, forthcoming).

important, such funding schemes may subvert the very ability of a private sphere to provide a countermeasure to government power.

*Rust* illustrates the way in which government funding may have both more potent and more disruptive effects than direct government speech, even holding expenditures constant. The impact of the government's own speech on abortion questions likely pales in comparison to the impact of advice and counseling given to pregnant women by health care providers. (The reason relates not only to the source of the speech—an apparently independent professional—but also to the time at which it occurs.) How better, then, to communicate an anti-abortion message: through direct speech or through selective subsidization of health care providers? The latter course amplifies the government's own message at the same time as (and partly because) it wrecks havoc on the ability of those private parties in the best position to challenge the message to provide a counterweight to government authority.<sup>70</sup>

But even if, or to the extent that, government funding decisions can be equated with government speech, so too can other content-based underinclusive government actions. Suppose (to borrow a hypothetical from Justice Scalia's opinion for the Court in *R.A.V.*) a city council enacts an ordinance prohibiting those legally obscene works—but only those legally obscene works—that do not include an endorsement of the municipal government.<sup>71</sup> The hypothetical involves an exemption from otherwise permissible regulation, rather than a direct cash grant or an exemption from taxation. Yet, as shown previously, no reason exists for treating the one as different from the others. In the regulatory exemption case, the government is still paying for speech in every significant respect: the speaker receives a benefit for expressing views supportive of city government, and the government absorbs costs of the expression that normally would be borne by the speaker. The mechanism is different, but the essential act is the same. If the government

---

<sup>70</sup> I do not claim that every government funding program will pose these dangers or that no funding program should be assimilated to government speech. A funding program may be constructed in so narrow a fashion as to appear identical (or nearly so) to the government's own expression. This will be true when the constitutional concerns I have discussed are slight or absent. But as I will show, the same may be said of other (non-funding) decisions involving content-based underinclusion. The fact of funding is neither necessary nor sufficient to transform content-based underinclusive action into government expression.

<sup>71</sup> 112 S Ct at 2543.

“speaks” when it pays for speech by private parties, then the government is speaking in the *R.A.V.* Court’s hypothetical.

The point can be made across the entire range of content-based underinclusion cases. In *Rust*, of course, the government made a direct cash grant for some kinds of expression, but not for others. In *R.A.V.*, which Justice Scalia saw as perfectly analogous to his obscenity hypothetical, the government offered some expression an exemption from otherwise applicable regulation of a proscribable speech category. The same mechanism is involved in cases, such as *Metromedia*, in which certain kinds of speech receive an exemption from otherwise reasonable time, place, or manner restrictions on expressive activity. And in some sense, the non-public forum cases bridge the gap: a rule that allows certain speech but not other speech on, say, a military base, as in *Greer*, can be viewed either as a direct grant (of certain rights in property, rather than of cash) or as an exemption from a generally applicable regulation prohibiting speech in a certain context. The key point is that the government actions in all these cases stand in a similar relation to government speech: in all, the government uses its powers, within a sphere of general discretion, to pick up the costs of speech—to pay for speech—of a particular content.

The argument based on government speech thus appears of limited consequence. The argument does not successfully challenge my central thesis: that there exists a single category of content-based underinclusion cases, all of which—regardless whether they involve direct funding—raise the same First Amendment issue. Nor does the government speech approach provide a comprehensive way of dealing with this issue. We can doubtless find instances of content-based underinclusion—again, some involving direct funding, some not—in which the government appears to be doing little more than speaking itself.<sup>72</sup> Yet surely, with respect to each

---

<sup>72</sup> In the non-public forum context, for example, we might wonder about a legal doctrine that would permit a general to speak to troops on a restricted military base about, say, alcohol use, but would preclude the general from inviting an expert on alcohol dependency to give a similar speech. An example of this kind suggests that courts might well recognize the possibility that, in a particular case, speech by a nominally private party should be treated as government speech. The inquiry should focus on the concerns mentioned above: whether the speech is clearly attributable to the government and whether the government’s action, in promoting the speech, threatens to interfere with the realm of private discourse in a way direct government speech would not. Indeed, it is possible that even direct government expression should be tested by standards of a similar kind.

kind of content-based underinclusion mentioned, we will find many (almost certainly, many more) cases in which the government, through use of its discretionary funding or regulatory powers, is doing something more than speaking—is in fact influencing and shaping the world of private discourse in a way that accords with its own beliefs of what kinds of speech should be promoted. *R.A.V.* arguably is one example; *Rust* arguably is another. To treat all this as permissible government speech is to ignore the scope and effect of the government action and the constitutional problems such actions may raise. It is to evade the critical question: In a sphere of general discretion over speech, when may government prefer private speech of a certain content to private speech of another?

#### IV

The cases I have discussed raise a common First Amendment issue and call for a common constitutional analysis. I do not suggest that all cases of content-based underinclusion must “come out” in the same manner. I do not, for example, assert that if *R.A.V.* is right, then *Rust* must be wrong, or vice versa. I claim only that these cases, and others raising the issue of content-based underinclusion, should be subjected to the same constitutional standards.

Establishing those standards is no easy task. The problem of selective funding alone has confounded generations of judges and constitutional scholars. I have argued that selective funding cases must be assimilated to other instances of content-based underinclusion. The difficulty, therefore, far from being eased, is in fact broadened.

In this part, I thus offer a preliminary—and necessarily sketchy—view of the proper constitutional approach to cases raising the issue of content-based underinclusion. I start by sorting through, in a more concrete fashion than I have done before, the diverse and conflicting ways the Court has responded to this problem. I then suggest, taking into account the effect and motive of government action, a distinction between two kinds of content-based underinclusion: that involving subject matter, which generally is acceptable; and that involving viewpoint, which generally is not. Finally, harking back to *Rust* and especially to *R.A.V.*, I pro-

pose certain modifications to this simple division of the cases—instances in which subject matter-based distinctions should raise constitutional concern and, perhaps too, instances in which viewpoint-based distinctions should be tolerated.

The Court, failing to recognize the common problem of content-based underinclusion, has employed a variety of constitutional standards in the kinds of cases discussed in this article. At one extreme, the Court has indicated that within a sphere of general discretion, the government has near-complete freedom to make content-based distinctions with respect to speech. At the other extreme, the Court has stated that the government is barred (at least in the absence of the most compelling justification) from making any such distinctions. Between these two positions lie others, sometimes only half-articulated, premised on the notion that not all content-based distinctions are alike. Thus, the Court at times has indicated that within an area of general discretion, the government may restrict speech on the basis of subject matter or speaker, but not on the basis of viewpoint. These various standards sometimes correspond to the different contexts in which the problem of content-based underinclusion arises, so that in each context a single standard holds sway. More confusingly, a plurality of these standards may coexist and compete within even a single subcategory of content-based underinclusion cases.

The greatest disarray, as I have noted, appears in the selective funding cases, in which the Court has adopted the full range of positions just described. Prior to *Rust*, the Court had indicated that in the funding context, some kinds of content discrimination mattered profoundly, though precisely what kinds remained uncertain. Thus, in *Arkansas Writers' Project, Inc. v Ragland*,<sup>73</sup> the Court explicitly rejected any distinction between subject matter-based and viewpoint-based regulation, stating that all content-based regulation was subject to strict scrutiny.<sup>74</sup> By contrast, in *Regan v Taxation with Representation*,<sup>75</sup> the Court held that the government, in

---

<sup>73</sup> 481 US 221 (1987).

<sup>74</sup> *Id.* at 230. The stringency of the Court's analysis may be attributable to a special concern about press regulation. The Court emphasized that "selective taxation of the press—either singling out the press as a whole or targeting individual members of the press—poses a particular danger of abuse by the State." *Id.* at 228. A standard so strict applying to all funding decisions would prevent almost all government funding of expression.

<sup>75</sup> 461 US 540 (1983).

funding speech, could make some kinds of content-based distinctions, but suggested in dicta that funding on the basis of viewpoint would violate the Constitution.<sup>76</sup> Finally, in *Rust* the Court took the position that the government could fund expression as it wished, in accordance with its "value judgments."<sup>77</sup> In the context of funding, the whole question of content discrimination—including viewpoint discrimination—became irrelevant.

In each of the other contexts discussed in this article, the Court has concluded that even within a sphere of general discretion, the First Amendment prohibits the government from making certain kinds of content distinctions; the Court, however, has adopted a less rigorous approach in non-public forum cases than in others. In the non-public forum cases, the Court has denied the government only the power to make viewpoint distinctions; regulations based on subject matter or speaker identity, so long as they satisfy a toothless reasonableness inquiry, are permitted.<sup>78</sup> By contrast, in cases such as *Metromedia* or *Mosley*, in which the Court considered limited time, place, or manner regulations involving either no public property or a public forum, the Court generally has applied strict scrutiny to all content-based exemptions, regardless whether the exemptions pertain to particular viewpoints or to more general subject matter categories. Here, the Court repeatedly has held that the government "may not choose the appropriate subjects for public discourse," even if, in doing so, "the government does not favor one side over another."<sup>79</sup>

The Court in *R.A.V.* leaned toward the position taken in cases such as *Mosley*, although with numerous hedges and qualifications.

---

<sup>76</sup> Id at 548, 550 (disapproving funding decisions "'aimed at the suppression of dangerous ideas'" (quoting *Cammarano v United States*, 358 US 498 (1959)); id at 551 ("[A] statute designed to discourage the expression of particular views would present a very different question.") (Blackmun concurring). The Court, in approving speaker-based funding decisions and disapproving viewpoint-based funding decisions, expressed no opinion on the permissibility of funding decisions based on the subject matter of speech. In other cases, however, the Court has treated similarly speaker-based and subject matter-based restrictions, distinguishing both from restrictions based on viewpoint. See, for example, *Perry v Perry*, 460 US 37 (1983); *Cornelius v NAACP*, 473 US 788 (1985).

<sup>77</sup> 111 S Ct at 1772.

<sup>78</sup> Thus, for example, the Court in *Greer v Spock*, 424 US 828 (1976), allowed a military base to exclude all partisan political speakers, and the Court in *Lebman v City of Baker Heights*, 418 US 298 (1974), permitted a municipal transportation system to refuse to post political advertisements. See also *Cornelius*, 473 US at 806; *Perry*, 460 US at 49.

<sup>79</sup> *Metromedia, Inc. v San Diego*, 453 US 490, 515, 518 (1981) (plurality); see *Carey v Brown*, 447 US 455, 460–61, 462 n 6 (1980); *Police Dep't v Mosley*, 408 US 92, 95, 99 (1972).

The *R.A.V.* Court, of course, ruled that at least some content-based distinctions within a proscribable category of speech violate the Constitution: “the First Amendment imposes . . . a ‘content discrimination’ limitation upon a State’s prohibition of proscribable speech.”<sup>80</sup> But what is the exact content of this limitation? The Court made clear that in the context of proscribable speech, the constitutional ban extends beyond explicit viewpoint-based distinctions; indeed, in the first statement of its holding, the Court declared the St. Paul law unconstitutional because it made distinctions “solely on the basis of the subjects the speech addresses.”<sup>81</sup> Yet the Court declined to say that in this sphere the First Amendment renders suspect all content-based restrictions: “the prohibition against content discrimination is not absolute. It applies differently in the context of proscribable speech than in the area of fully protected speech.”<sup>82</sup> Repeatedly asking whether a regulation would pose a “significant danger of idea or viewpoint discrimination,” the Court listed a series of constitutionally unobjectionable content-based distinctions.<sup>83</sup> The list closed with the suggestion that, within a proscribable category of speech, content-based distinctions may be permissible so long as they present “no realistic possi-

---

<sup>80</sup> 112 S Ct at 2545.

<sup>81</sup> Id at 2542. The Court later concluded that the ordinance also discriminated with regard to viewpoint, but as I will discuss, this argument at least raised questions; the Court’s decision thus depended heavily on the ban on subject matter restrictions. With respect to this ban, the majority opinion differed not only from Justice White’s approach, but also from Justice Stevens’s alternative analysis. Unlike Justice White, Justice Stevens would view certain content-based distinctions within proscribable categories of speech as constitutionally troubling. But Justice Stevens, unlike the *R.A.V.* majority, apparently would accord automatic strict scrutiny only to those content distinctions based explicitly on viewpoint. See id at 2568–69.

<sup>82</sup> Id at 2545.

<sup>83</sup> Id at 2545–47. First on the list were distinctions supported by the very factor that rendered the entire category of speech proscribable. To use one of Justice Scalia’s examples, the government could prohibit, from the broad category of legally obscene materials, only the “most lascivious displays of sexual activity.” Id at 2546. As each of the concurrences noted, this exception may have covered the St. Paul ordinance, which reasonably could be viewed as an attempt to prohibit, from the entire category of fighting words, those which “by their very utterance” inflict the greatest injury or pose the greatest danger of retaliatory violence. See id at 2556, 2565. Justice Scalia also excepted from rigorous constitutional scrutiny laws containing content distinctions based on the “secondary effects” (i.e., noncommunicative effects) of speech, as well as laws directed against conduct but incidentally covering a content-based subcategory of proscribable speech. See id at 2546–47. Finally, Justice Scalia would have viewed more leniently (although his reasoning on this count is mysterious) a prohibition of speech falling within a proscribable category that is “directed at certain persons or groups,” id at 2548—yet another exception that reasonably could have been used to insulate the St. Paul ordinance from strict review.

bility that official suppression of ideas is afoot."<sup>84</sup> Whether a regulation prohibiting expression on certain subjects ever could fall within this "general exception" to the ban on content discrimination was left uncertain.

What then is the right approach? When, if ever, will some manner of content-based underinclusion invalidate a speech regulation? As I have said, the same constitutional standards should govern all of the various kinds of cases discussed in this article. I do not mean to suggest that the government interests underlying the underinclusive regulation of speech will be identical in all contexts. The nature of the government action at issue—for example, direct funding of speech or regulation of speech within a non-public forum—will sometimes provide distinctive justifications for content-based underinclusion.<sup>85</sup> Thus, in acting as manager of a military base, the government may have—as it claimed to have in *Greer*—peculiar reasons for restricting some speech, such as the interest in insulating a military establishment from partisan political causes. Similarly, in providing direct funding out of public coffers, the government frequently will have to take into account the limited availability of revenues devoted to a particular program or purpose. But because each kind of government action discussed in this article affects First Amendment rights in the same way, each should be held to the same set of justificatory burdens. The remaining question concerns the appropriate content of these burdens. That question is best approached by focusing on the nature of the First Amendment problem in all of these cases.

Thus recall what the Court confronts in each one of these contexts. The government is operating within a sphere of general discretion: it can refuse to promote or allow any speech at all. Instead, the government chooses to advance or permit some, but not other, speech on the basis of content. If the Court strikes down the action, citing content discrimination, the government can return to a general ban, becoming (in terms of total quantity of speech) more, rather than less, speech restrictive. The government can prohibit all fighting words, can bar all speakers from a military base, can

---

<sup>84</sup> *Id.* at 2547. As an illustration of a content-based distinction posing no threat of censorship of ideas, Justice Scalia hypothesized an ordinance prohibiting only those obscene motion pictures featuring blue-eyed actresses.

<sup>85</sup> *Cf.* Sullivan, 102 Harv L Rev at 1503 (cited in note 58); Sunstein, *Why the Unconstitutional Conditions Doctrine Is an Anachronism*, 70 BU L Rev at 607 (cited in note 58).

prevent any person from using a noisy soundtrack, can decline to fund any speech. If all this is so, one way to approach the problem at least becomes clear. What we need to ask is when content discrimination resulting in more speech is of greater constitutional concern than content neutrality resulting in less. We can begin, in other words, to tackle the essential issue in all of these cases by rephrasing it (somewhat crudely) in the following terms: When is some speech worse than none?<sup>86</sup>

A proper response to this inquiry should focus on both the effects and the purposes of content-based underinclusive action. In other words, government regulation allowing some speech may raise greater constitutional problems than regulation allowing no speech at all either because the former has graver consequences than the latter or because the former more likely proceeds from an improper impulse. Both considerations suggest an initial, broad distinction between underinclusive action based on viewpoint and underinclusive action based on subject matter.

Consider first the possible consequences of underinclusive regulation of speech on the realm of public discourse.<sup>87</sup> Sometimes, such regulation will place particular messages at a comparative disadvantage and, in doing so, will distort public debate. An example is Justice Scalia's hypothetical ordinance prohibiting all legally obscene materials except those containing an endorsement of city government. Such a law leaves untouched speech supportive of city government, while restricting speech critical of city government, thereby skewing discourse on this issue. That obscenity (like fight-

---

<sup>86</sup> It might be argued that framing the inquiry in this way assumes unjustifiably that the government will respond to the invalidation of a content-based distinction by expanding the reach of the speech restriction, rather than by eliminating it entirely. This objection recognizes, quite correctly, that in some circumstances an apparently "greater" power is in fact practically or politically constrained; in that event, if the "lesser" power is removed, the government will not exercise its authority at all. See Kreimer, 132 U Pa L Rev at 1313 (cited in note 58). But in the settings discussed in this article, the objection appears to have only slight weight. The more expansive powers here—enacting limited time, place, or manner restrictions, establishing broad speech restrictions for non-public forums, declining to fund speech, proscribing categories of speech like fighting words or obscenity—are in most instances not merely theoretically but actually available; the government very frequently exercises such powers. We indeed may wish to keep in mind that in some cases, the government as a practical matter will not be able to—or, perhaps more frequently, will not wish to—expand the coverage of a speech restriction, but the central inquiry in these contexts remains as I have described it in the text.

<sup>87</sup> See Stone, *Content Regulation and the First Amendment*, 25 Wm & Mary L Rev at 198–200, 217–27 (cited in note 40), for a full discussion of these issues in connection not with content-based underinclusion, but with content-based discrimination generally.

ing words) is by definition unprotected makes no difference to the analysis; the distortion relates to ideas and messages extrinsic to that category. It is true that the distorting effect occurs at the margin; persons opposed to city government can communicate this message through means other than obscenity. Yet the ordinance remains more constitutionally problematic than a total ban on obscenity, which would have no skewing effect at all on the debate concerning city government.<sup>88</sup> Precisely the same point can be made in the context of direct funding. Assume our city council, informed of the decisions in *R.A.V.* and *Rust*, instead passed a law providing for public funding of all speech endorsing incumbent city officials in their campaigns for reelection. Such a law similarly provides a comparative advantage to messages of endorsement, thereby again skewing public debate. As with the obscenity statute, the skewing effect makes the statute more troublesome than a complete absence of public funding.<sup>89</sup>

Not all instances of content-based underinclusion, however, will have such problematic effects. Contrast to the viewpoint-based laws used above a set of regulations discriminating in terms of general subject matter. First, suppose that the city council enacts a law prohibiting all obscene materials except those dealing in any way with government affairs. It is no longer so clear that a total ban on obscenity would better serve First Amendment interests. At least facially, the law does not skew public debate about matters

---

<sup>88</sup> Of course, a total ban on obscenity removes all obscene messages from the world of public discourse, which in some other world might be thought a constitutional problem of large dimension. The premise here—accepted by the Supreme Court—is that eliminating obscenity per se from the realm of public debate raises no First Amendment problem whatsoever. A premise of similar kind exists in all cases of content-based underinclusion.

<sup>89</sup> The notion of a skewing effect, as set forth in the text, of course assumes that distortion arises from government, rather than from private, action. That assumption may be misplaced. If there is "too much" expression of a particular idea in an unregulated world, then government action specially disfavoring that idea might "un-skew," rather than skew, public discourse. See Fiss, 100 Harv L Rev at 786–87 (cited in note 53); Sunstein, *Free Speech Now*, 59 U Chi L Rev at 295–97 (cited in note 53). An understanding of this point has special relevance in considering underinclusive government action. With respect to such actions, the only constitutional worry is equality among ideas; restriction, taken alone, need not concern us. The situation is very different in the case of other kinds of speech restrictions, whose unconstitutionality may rest as much or more in considerations of personal autonomy as in considerations of equality. See generally David A. Strauss, *Persuasion, Autonomy, and Freedom of Expression*, 91 Colum L Rev 334 (1991). Nonetheless, I think the assumption used here to measure distortion is generally, although not invariably, proper. Any other would allow the government too great—and too dangerous—an authority to decide what ideas are overrepresented or underrepresented in the market.

involving government, as the viewpoint-based obscenity ordinance did.<sup>90</sup> Of course, the law allows the use of obscene materials to speak about government affairs, while restricting the use of those materials to speak about a host of other subjects. But neither those who wish to speak on such subjects nor their potential audience can claim in any real sense that the ordinance harms them more than would a ban on all obscene materials. The law, viewed solely in terms of effects on public debate, thus appears consistent with the First Amendment. And once again, the same is true of a similar statute involving the mechanism of direct funding. Assume that the city council passes a law providing for public funding of all candidates for elected office. Here too, the statute makes a content-based distinction: one kind of speech is funded, all other speech is not. But as long as the law covers all candidates and parties, no one can complain that the subsidy plan has effects on public debate that are constitutionally more troublesome than a refusal to subsidize at all.<sup>91</sup>

Yet effects are not all that matter in considering the permissibility of content-based underinclusion; we also must take into account the purposes underlying the government action.<sup>92</sup> Notwithstanding that another, more speech restrictive action could have been taken (assuming a proper purpose), the purpose of *this* action—the action in fact taken—must fall within the range of constitutional legitimacy. What objectives fall outside that range? It is a staple of First Amendment law that no government action may be taken because public officials disapprove of the message communicated. The flip side of this principle, as Geoffrey Stone has noted, is that “the government may not exempt expression from an otherwise general restriction because it agrees with the speaker’s views.”<sup>93</sup> Thus, as the *R.A.V.* Court stated: “The government may not regulate use [of fighting words] based on hostility—or favoritism—towards the

---

<sup>90</sup> I consider at text accompanying note 110 problems relating to viewpoint-differential consequences of such facially viewpoint-neutral laws. It may well be that this statute looks sufficiently odd to heighten concerns about such consequences.

<sup>91</sup> In covering all parties and candidates, the hypothetical statute stands on firmer ground than the subsidy scheme approved in *Buckley v Valeo*, 424 US 1 (1976), which funds some candidates and not others and thus may well distort debate on critical public matters.

<sup>92</sup> Again, Geoffrey Stone provides a fuller discussion of these issues, in the context of discussing content-based discrimination generally, in *Content Regulation and the First Amendment*, 25 Wm & Mary L Rev at 212–17, 227–33 (cited in note 40).

<sup>93</sup> *Id.* at 228.

underlying message expressed."<sup>94</sup> Other constitutionally disfavored justifications for government action also appear in the cases—most notably, that the government may not restrict expression because it will offend others. Once again, as said in *R.A.V.*, selective limitations on speech may not be justified by "majority preferences."<sup>95</sup> Regardless whether the government could achieve the same or greater effects with another end in mind, the existence of such illegitimate aims should invalidate the action at issue.

The distinction between viewpoint-based restrictions and subject matter-based restrictions serves as a useful proxy in evaluating the purpose, as in evaluating the effects, of underinclusion. A return to the set of hypotheticals offered above illustrates this point. The actions singling out for favorable treatment endorsements of city government can be presumed to stem from an illegitimate motive: what legitimate reason could lie behind these regulations? A similar danger presents itself with regard to any government action favoring or disfavoring a particular viewpoint: if suppression of the viewpoint does not lie directly behind the action, at least attitudes toward the viewpoint may influence the decision.<sup>96</sup> By contrast, government actions covering speech of a variety of viewpoints, even if on a single topic, less probably emerge from government (or majority) approval or disapproval of a particular message, precisely because they apply to a range of diverse messages. So, for example, the statute providing funds for campaign speech likely stems from a desire to reduce corruption, and the ordinance granting an exemption to obscenity involving discussion of government affairs may arise from the view (common and usually permissible in First Amendment law, though reflecting a kind of favoritism) that political speech is of special constitutional value.<sup>97</sup> The key point is that just as subject matter restrictions will less often skew debate than viewpoint restrictions, so too will they less often arise from constitutionally improper justifications.<sup>98</sup>

---

<sup>94</sup> 112 S Ct at 2545.

<sup>95</sup> *Id.* at 2548.

<sup>96</sup> See Stone, *Content Regulation and the First Amendment*, 25 *Wm & Mary L Rev* at 231 (cited in note 40).

<sup>97</sup> Again, however, this hypothetical regulation seems so eccentric that a closer examination into both purpose and effects might be in order. See note 90 and text at note 110.

<sup>98</sup> See Geoffrey R. Stone, *Restrictions of Speech Because of Its Content: The Peculiar Case of Subject-Matter Restrictions*, 46 *U Chi L Rev* 81, 108 (1978).

So far, then, we appear to have a simple way to test government action of the kind this article addresses. Viewpoint-based regulation should receive the strictest constitutional scrutiny, both because it skews public debate in a way a general ban (or refusal to subsidize) would not and because it more likely arises from an impermissible motive. By contrast, subject matter-based regulation, which generally raises concerns of purpose and effect no greater than would a general ban, should receive less searching examination, involving (as in the case of content-neutral regulations) a general balancing analysis.

Thus, for example, in *Rust*, the Court first would decide whether the selective subsidization rested on the speaker's viewpoint. There seems little serious argument on this score: the regulations, quite explicitly, prohibited funded projects from "encourag[ing], promot[ing] or advocat[ing] abortion," as well as from engaging in abortion referral and counseling; at the same time, the regulations permitted funded projects to engage in anti-abortion advocacy and required them to refer women for prenatal care and adoption services.<sup>99</sup> Once the determination of viewpoint discrimination is made in this manner, a strong presumption of unconstitutionality would attach, rebuttable only upon a showing of great need and near-perfect fit. If the government could not make this showing, the subsidization scheme would be struck down, leaving the government with the option of funding either less or more speech relating to abortion.

This result accords with the principles, relating to the purpose and effects of government regulation, underlying a strict presumption against viewpoint-based underinclusion. The regulations at issue in *Rust* can hardly be understood except as stemming from government hostility toward some ideas (and their consequences) and government approval of others: the subsidization scheme, as the majority itself noted, reflected and incorporated a "value judgment."<sup>100</sup> Further, the regulations, in treating differently opposing points of view on a single public debate, benefitted some ideas at the direct expense of others and thereby tilted the debate to one side. For both these reasons, a refusal to fund any speech relating to

<sup>99</sup> 42 CFR §§ 59.8(a)(2), 59.8(b)(4), 59.10, 59.10(a) (1990); 53 Fed Reg 2927 (1988).

<sup>100</sup> 111 S Ct at 1772.

abortion would have been constitutionally preferable to the funding scheme that the regulations established.

Before this analysis becomes too comfortable, however, a final look at *R.A.V.* is in order. That case, far more than *Rust*, poses serious challenges—on every level—to the simple approach suggested so far: to the ability to distinguish between viewpoint-based and subject matter-based underinclusion, to the relaxed constitutional standard applying to subject matter-based underinclusion, and to the presumed impermissibility of viewpoint-based underinclusion. In so doing, *R.A.V.* forces modifications to the analytical structure presented thus far, as well as a continued willingness to test that structure against the concerns of purpose and effect giving rise to it.

To see the difficulties *R.A.V.* presents, we should consider, as an initial matter, whether the St. Paul ordinance discriminated on the basis of viewpoint or subject matter. This undertaking involves three separate inquiries: first, whether the ordinance on its face discriminated on the basis of viewpoint or subject matter; second; whether the ordinance in practice discriminated on the basis of viewpoint or subject matter; and third, which measure of discrimination (facial or operational) is to control if the answers to the first two questions differ. In exploring these issues, and attempting to draw more general lessons from them, I will refer frequently to Justice Scalia's and Justice Stevens's contrasting characterizations of the St. Paul ordinance.

Viewed purely on its face, the St. Paul ordinance, as construed by the Minnesota Supreme Court, appears to discriminate only on the basis of subject matter. The ordinance proscribed such fighting words as caused injury on the basis of race, color, creed, religion, or gender—that is, such fighting words as caused injury on the basis of certain selected topics. For this reason, Justice Stevens viewed the ordinance as at most a subject matter restriction:<sup>101</sup> all fighting words, uttered by any speaker of whatever viewpoint, concerning another person's "race, color, creed, religion, or gender" were forbidden. Even Justice Scalia frequently referred to the ordinance in this manner; in apparent acknowledgment of the

---

<sup>101</sup> Justice Stevens initially argued that the ordinance was based neither on viewpoint nor on subject matter, but only on the injury caused by the expression. 112 S Ct at 2570. For discussion of this point, see text at notes 116–17.

statutory language, he described the law as regulating expression “addressed to . . . specified disfavored topics,” as policing “disfavored subjects,” and as “prohibit[ing] . . . speech solely on the basis of the subjects the speech addresses.”<sup>102</sup> Thus, if the analysis I have proposed is correct, and if a law is to be classified as viewpoint based or subject matter based solely by looking to the face of the statute, then Justice Scalia erred in finding the discrimination worked by the statute to be unconstitutional.

Beyond the question of facial discrimination, however, lurked another issue: Did the statute discriminate in its operation on the basis of viewpoint? Justice Stevens insisted that it did not. Describing how the ordinance would apply to both sides of a disputed issue, Justice Stevens noted: “[J]ust as the ordinance would prohibit a Muslim from hoisting a sign claiming that all Catholics were misbegotten, so the ordinance would bar a Catholic from hoisting a similar sign attacking Muslims.”<sup>103</sup> Or (to take a simpler example) just as the ordinance would prevent the use of racial slurs by whites against blacks, so too would it prevent the use of racial slurs by blacks against whites.<sup>104</sup> Justice Scalia admitted this much, but nonetheless suggested that the ordinance operated in a viewpoint discriminatory manner. In some debates, Justice Scalia reasoned, the regulation would “license one side of a debate to fight freestyle, while requiring the other to follow Marquis of Queensbury Rules.”<sup>105</sup> As an example, Justice Scalia noted that a sign saying that all Catholics were misbegotten would be prohibited, because the sign would insult on the basis of religion, but a sign saying that all anti-Catholic bigots were misbegotten would be permitted.

The conflict between Justice Scalia and Justice Stevens on this point serves as a reminder that the decision whether a statute dis-

---

<sup>102</sup> 112 S Ct at 2542, 2547; see *id.* at 2570 (Stevens dissenting).

<sup>103</sup> 112 S Ct at 2571. Justice Stevens assumed in this example that the signs would constitute fighting words.

<sup>104</sup> Akhil Amar makes the interesting point that Justice Stevens seemed to go out of his way to avoid this obvious example, using instead a hypothetical involving two minority groups. Amar notes too that Justice White’s opinion appeared to assume that the statute was asymmetrical, in the sense that it protected vulnerable social groups from dominant social groups, but not vice versa. See Amar, 106 Harv L Rev at 148–50 (cited in note 24). To the extent the statute is read in this manner—and Amar points out that the explicit examples in the statute (burning crosses and swastikas) are consistent with this reading—the viewpoint discrimination inherent in the statute becomes quite obvious.

<sup>105</sup> *Id.* at 2548.

criminate on the basis of viewpoint may be highly contestable.<sup>106</sup> The very notion of viewpoint discrimination rests on a background understanding of a disputed issue. If one sees no dispute, one will see no viewpoints, and correspondingly one will see no viewpoint discrimination in any action the government takes.<sup>107</sup> Similarly, how one defines a dispute will have an effect on whether one sees a government action as viewpoint discriminatory. Justice Stevens understood the public debate on which the St. Paul ordinance acted as a dispute between racism of different stripes.<sup>108</sup> With respect to this dispute, the ordinance took a neutral position and effected a neutral result. Justice Scalia, by contrast, saw the dispute as one between racists and their targets and/or opponents. With respect to this dispute, the ordinance appeared to take a side. By prohibiting fighting words based on race, while allowing other fighting words, the law barred only the fighting words that the racists (and not the fighting words that their targets) would wish to use.

In this conflict, Justice Scalia seems to me to have the upper hand: the St. Paul ordinance, in operation, indeed effected a form of viewpoint discrimination. We can all agree that a law applies in a viewpoint discriminatory manner when it takes one side of a public debate. We should also all be able to agree that one way of taking sides is by handicapping a single contestant—and further, that one way of handicapping a contestant is by denying her a particular means of communication (such as fighting words).<sup>109</sup> The

<sup>106</sup> The difficulty may arise in considering either facial or operational viewpoint discrimination. Had the ordinance, on its face, prohibited all racist fighting words, the debate between Justice Scalia and Justice Stevens presumably would have been the same. Justice Stevens would have argued that the statute on its face did not discriminate on the basis of viewpoint because it prohibited all kinds of racist fighting words. Justice Scalia, by contrast, would have argued that the statute was facially viewpoint discriminatory because it prohibited the fighting words used by racists, but not the fighting words directed at them.

<sup>107</sup> See Catharine A. MacKinnon, *Feminism Unmodified* (Harvard, 1987) at 212 (“What is and is not a viewpoint, much less a prohibited one, is a matter of individual values and social consensus.”).

<sup>108</sup> Justice Stevens at one point acknowledges a debate between proponents of bigotry and proponents of tolerance, but he insists that the ordinance also is neutral with respect to this debate. Thus, Justice Stevens says that the “response to a sign saying that ‘all [religious] bigots are misbegotten’ is a sign saying that ‘all advocates of religious tolerance’ are misbegotten.” 112 S Ct at 2571. This statement has a lovely symmetry, but also a sense of unreality. Presumably, bigots wish to direct their speech not to abstract advocates of tolerance, but to members of a despised group. The question *R.A.V.* presents is whether the government can impose limits on the bigots’ desire to do so. Here, Justice Stevens ignores this issue by reframing the public debate.

<sup>109</sup> That a regulation deprives a speaker only of a particular means of communication does not make the regulation any less an example of viewpoint discrimination. Indeed, almost all

St. Paul ordinance, it is true, handicaps both sides (and therefore neither side) when Jews and Catholics, whites and blacks scream slurs based on religion or race at each other. But surely race-based fighting words occur (indeed, surely they usually occur) in something other than this double-barreled context. In most instances, race-based fighting words will be all on one side, because only racists use race-based fighting words, and racists usually do not assail only each other. When the dispute is of this kind, the government effectively favors a side in barring only race-based fighting words. To put the point another way, if a law prohibiting the display of swastikas takes a side, no less does a law that punishes as well the burning of crosses.

Yet even if this is so, the question remains how to categorize a statute (such as the St. Paul ordinance) that discriminates on the basis of viewpoint only in operation, and not on its face. Do we classify the St. Paul ordinance as a subject matter restriction (in keeping with the face of the statute) or as a viewpoint restriction (in keeping with the way it works in practice)? Or, to put the question in a more meaningful way, regardless of the label we attach to the statute, do we treat it as discriminating on the basis of viewpoint or of subject matter?

When a statute has so unbalanced a practical effect as the St. Paul ordinance, I think, it must be treated in much the same manner as a statute that makes viewpoint distinctions on its face. I have argued that underinclusive actions based on subject matter generally should receive relaxed scrutiny because they pose little danger of skewing public debate on an issue or arising from an illegitimate motive; thus, they usually will be no worse (and because less speech restrictive, often a great deal better) than a refusal to allow or subsidize any speech at all. But a subject matter restriction of the kind in *R.A.V.* flouts this reasoning. Here, the restriction, although phrased in terms of subject matter, meaningfully applied only to one side of a debate and thus had a tilting effect as profound as a

---

cases of underinclusion function only to remove a particular means of communication from the speaker: the speaker may not use fighting words; the speaker may not use a noisy soundtruck; the speaker may not use the grounds of a military base; the speaker may not use government funds. In all of these cases, the government does not act to eliminate completely an idea from the realm of public discourse, but may nonetheless take a side. That the government's action deprives a speaker only of a means of communication is relevant, if at all, not to the question whether the action is viewpoint-based, but to the question whether, even if viewpoint-based, the action should be allowed.

viewpoint-based regulation; the ordinance, though facially prohibiting "race-based" fighting words, might as well have prohibited racist fighting words—that is, fighting words expressing the view of racism. And precisely because the law operated in this way, the likelihood that it stemmed from impermissible motives must be treated seriously; knowing that the ordinance would restrict only a particular point of view, legislators might well have let their own opinion, or the majority's opinion, of that viewpoint influence their voting decision.<sup>110</sup> The ordinance thus presented the same dangers as a facially viewpoint-based speech regulation.

It might be argued that in admitting this much, I have compromised fatally the position that underinclusive actions based on subject matter generally should not be subject to strict constitutional scrutiny. After all, many subject matter restrictions have viewpoint-differential effects; in all such cases, it might be said, precisely the same arguments for strict scrutiny would apply. Further, the argument might run, it may be difficult to distinguish these subject matter restrictions from others, and it may be wise as a general matter to overprotect speech; thus, we perhaps should look upon all subject matter restrictions with suspicion. But this argument ignores the special feature of underinclusion cases: that in such cases, invalidating a subject matter restriction will as likely (perhaps more likely) lead to less, as to more, expression. In this kind of case, a defensive, overprotective approach seems inappropriate: we should treat subject matter restrictions harshly only when they pose real dangers of distorting effects or impermissible motive. To the extent, then, that the *R.A.V.* opinion stands for the proposition that all content-based underinclusion violates the Constitution,<sup>111</sup> the opinion is in error.

This aspect of the analysis, no doubt, raises difficult questions. One set involves the determination at what point the viewpoint differential effects of a regulation that on its face involves subject matter alone should begin to give rise to suspicion. Need we worry only about statutes such as that involved in *R.A.V.*, in which the

---

<sup>110</sup> As the *R.A.V.* Court noted, St. Paul argued that the law was necessary, among other reasons, to show that speech expressing hatred of groups was "not condoned by the majority." 112 S Ct at 2548. It is difficult to conceive of a more illegitimate purpose for regulating speech.

<sup>111</sup> See text at notes 80–84 for discussion of the ambiguity of the *R.A.V.* opinion on this question.

regulation effectively restricts one side alone, or need we worry too about statutes with lesser, but still noticeable, viewpoint-based effects? Another set of questions involves the technique used to identify troublesome regulations. Should we use case-by-case analysis, or should we try to devise some more general standard to separate out the most dangerous restrictions based facially on subject matter? Whatever the precise answers to these questions, though, the basic point remains: on some occasions, a regulation that on its face involves only subject matter must be treated as if it involved viewpoint; on most occasions, it need not.

In this statement, however, a final question lurks: When, if ever, may we tolerate viewpoint-based underinclusive actions? Suppose, for example, that the government wished to fund private speech warning of the dangers of tobacco. Would the government also be required to fund private speech minimizing the health risks associated with smoking? One answer to this question is to insist on strict viewpoint neutrality in the support of private speech; then, if the government wished to express an anti-smoking message, it would have to disdain private speech and do the job itself. Yet this answer runs contrary to many of our intuitions. The same point can be made by using a hypothetical along the lines of *R.A.V.* Suppose that the government banned all (but only) those legally obscene materials that featured actors smoking cigarettes. Would this action seem any more objectionable than the example Justice Scalia gave of innocuous selectivity within a proscribable category—the prohibition of all (but only) those obscene materials featuring blue-eyed actresses?<sup>112</sup> The smoking ordinance may seem, if anything, less troublesome; it, at least, has a reason. And yet the ordinance discriminates on the basis of viewpoint.

I cannot here consider in detail the circumstances in which viewpoint-based underinclusion should be upheld. I will note, however, a few points that may serve to structure future inquiry regarding this issue. These relate, first, to the possibility that some viewpoint-based underinclusion may be adequately justified even under a compelling interest test, and, second, to the more remote possibility that some viewpoint-based underinclusion need not be subjected at all to this most stringent standard.

---

<sup>112</sup> 112 S Ct at 2547.

The initial point is—or should be—obvious: strict scrutiny need not invalidate a viewpoint-based underinclusive action. The test, as stated by the Court, is whether the regulation is both necessary and narrowly tailored to serve a compelling interest.<sup>113</sup> In *R.A.V.*, the Court mistakenly interpreted this test to create a *per se* rule against viewpoint underinclusion. Action of this kind, the Court said, is never necessary, because the government can always enact a broader speech regulation.<sup>114</sup> But if the speech additionally covered by a broad regulation fails to advance the interest asserted, why must the government restrict it as well? Assume, for example, that the government has a compelling interest in ensuring that children do not start smoking; assume as well that speech extolling cigarettes in the immediate vicinity of a school leads children to start smoking. Must the government, to prevent this speech, enact a law that restricts speech in the vicinity of schools to the full extent allowed under the Constitution? Would such a law be either “necessary” or “narrowly tailored” to serve the asserted interest? The questions answer themselves. A viewpoint-based underinclusive action should not be held invalid (as it was in *R.A.V.*) on the mere ground that it is, by definition, underinclusive. If the government can show—if, for example, St. Paul could have shown—that it has a compelling interest, that it must regulate speech to achieve that interest, and that it has regulated all (but only) such speech as is necessary to achieve the interest, then the government action should pass strict scrutiny.<sup>115</sup>

The second point I make more tentatively: indeed, I pose it as a question: Must all viewpoint-based underinclusive actions be subject to strict scrutiny, or are there some “viewpoints” that in the context of underinclusion need not be treated as such? The examples I have used, relating to viewpoints on tobacco use, seem to suggest that not all viewpoints are alike, although it is difficult to fashion a principled reason why. If our intuitions rebel against the idea that the government cannot fund speech discouraging

---

<sup>113</sup> See, for example, *Perry v Perry*, 460 US 37, 45 (1983); *Cornelius v NAACP*, 473 US 788, 800 (1985).

<sup>114</sup> See 112 S Ct at 2550.

<sup>115</sup> See *Burson v Freeman*, 112 S Ct 1846 (1992), for a recent First Amendment case in which the Court understood the compelling interest standard in this manner (although perhaps misapplied it). In keeping with the essential thesis of this article, I believe this standard should govern in all cases of viewpoint-based underinclusion, including funding decisions.

smoking without also funding its opposite, they do so for some combination of three reasons, each of which exists in tension with common First Amendment principles. First, the debate in this case, by its nature, offers the hope of right and wrong answers—answers subject to verification and proof. Second, society has reached a shared consensus on the issue; the answers, in addition to being verifiable, are widely believed. And third—and most important—one side of the debate appears to do great harm. When these factors join, a viewpoint regulation may appear justifiable whenever a more general regulation could exist. Then, government disapproval of a message may seem no longer illegitimate, because the disapproval emerges from demonstrable and acknowledged harms; then too, the distortion of debate resulting from the government action may appear not vice, but virtue. Some speech here seems better than none.

Justice Scalia's and Justice Stevens's opinions in *R.A.V.* included a debate on just these issues. Justice Stevens first characterized the St. Paul ordinance not as viewpoint-based, not even as subject matter-based, but as injury-based: the ordinance banned speech that caused a special and profound harm. Justice Scalia mocked this approach, dismissing it as "word-play": "What makes the [injury] produced by violation of this ordinance distinct from the [injury] produced by other fighting words is nothing other than the fact that it is caused by a distinctive idea, conveyed by a distinctive message."<sup>116</sup> Replied Justice Stevens: the Court failed to comprehend "the place of race in our social and political order"; were it to do so, it would recognize that race-based fighting words were a grave social evil, causing "qualitatively different" harms from other fighting words.<sup>117</sup> St. Paul, on this view, had done nothing more than respond, neutrally and legitimately, to real-life concerns; and any resulting skewing effect, given these concerns, need hardly trouble us. To put the position most starkly (more starkly than Justice Stevens did): Even if, in some technical sense, the statute involved viewpoint, it was viewpoint we could cease to recognize as such for purposes of constitutional analysis.

The position of Justice Stevens cannot be right as a general matter. Almost all viewpoint-based regulations can be viewed as "harm-based" regulations, responding neutrally not to ideas as

---

<sup>116</sup> 112 S Ct at 2548.

<sup>117</sup> *Id.* at 2565, 2570 n 9.

such, but to their practical consequences. We may indeed take as a given that almost all viewpoints anyone would wish to restrict cause arguable harms in some fashion. So, for example, in *Rust*, supporters of the regulations might argue that the selective funding corresponds not to viewpoints, but to demonstrable injuries (in the eyes of many) produced by abortion advocacy and counseling. And were we to treat such a case differently on the ground that there is no consensus on the "harmfulness" of this speech's consequences, then we would transform the First Amendment into its opposite—a safe haven for only accepted and conventional points of view.

Yet Justice Scalia's studied refusal to acknowledge or discuss the injuries caused by the speech in *R.A.V.* remains troubling. Here we have speech that, taken alone, has no claim to constitutional protection. The government responds to the special nature of this speech—to the special evil it causes—by in fact refusing to protect it. Perhaps this harm should be evaluated only in determining whether the government has met its high burden of justifying a distinction based on viewpoint. (Certainly, contrary to Justice Scalia's approach, the harm should be evaluated for this purpose.) The question that remains open for me is whether profound and indisputable harms can be taken into account for the purpose of lowering the standard of review applicable to viewpoint-based underinclusion—whether and when they may negate our usually justifiable concerns about the effects and motive of such government action. It may be possible to develop guidelines for this purpose—guidelines that will isolate and harshly confine a set of underinclusion cases in which viewpoint distinctions should be tolerated. But until we perform this feat, we could do far worse than to rely on a no-viewpoint distinction rule to handle cases of content-based underinclusion.

## V

For now, it may be less important to solve the problem of content-based underinclusion than to understand that there is a problem to be solved. My claim throughout this article has been that a certain set of cases—cases generally treated as if they have nothing in common with each other—raise a common issue and demand a common answer. The cases come in four general categories. The two most recently treated by the Court (though in widely

divergent ways) are typified by *Rust* and *R.A.V.*, the former involving selective funding of speech, the latter involving selective bans on speech within a wholly proscribable speech category. Add to these two others: cases involving selective bans on speech within a non-public forum and cases involving selective imposition of otherwise reasonable time, place, or manner restrictions, whether or not related to government property. The cases differ in context, but they share a structure transcending dissimilarities—a structure calling for acknowledgment by the Court and an effort to devise a uniform approach.

The problem these cases present is a problem of First Amendment neutrality, in as stark a form as can be found. In all these cases, the government may refuse to allow or subsidize any speech; the question remains when the government may refuse to allow or subsidize some (but not all) speech on the basis of content—when the government may give a special preference to expression of a certain kind. The cases cannot be distinguished by means of the subsidy/penalty distinction. The government action in all of these cases can be viewed as a subsidy; in each, the government voluntarily favors—and pays for—a certain kind of expression. More, labeling the action a subsidy or penalty is in these cases immaterial; assuming the government action constitutes a penalty, the problem lies not in the penalty itself, but in the government's selectivity—a problem that remains in the exact same form if the action is viewed a subsidy. For much the same reasons, the cases also cannot be distinguished by resort to an expansive notion of government speech. The action in all of these cases can be so characterized; and unless the government speech analogy has a power so far unsuspected in First Amendment law, it cannot displace the core issue in the cases. That issue must be confronted in whatever context it arises: when the government need not protect or promote any speech—when the speech itself has no claim upon the First Amendment—what limits remain on the government's power of selection?

I have suggested one approach to the problem; no doubt there are others worthy of attention. And were the Supreme Court to address the question in this way, no doubt the Justices would differ with respect to the solution. At least then, however, the debate in these cases would concern what under the First Amendment should matter. The answer might remain unclear, but the Court would have understood the question.

# Law & Social Inquiry

Journal of the American Bar Foundation  
Volume 18, Number 1, Winter 1993

A Libel Story: *Sullivan* Then and Now

Elena Kagan

The University of Chicago Press

---

# Law & Social Inquiry

---

Journal of the American Bar Foundation

Published by The University of Chicago Press

---

<i>Editors</i>	<b>Arthur F. McEvoy</b> <b>Elizabeth Mertz</b> <b>Peter Siegelman</b>
<i>Review Editor</i>	<b>Howard S. Erlanger</b>
<i>Associate Editors</i>	<b>Karyl A. Kinsey</b> <b>Christopher L. Tomlins</b>
<i>Managing Editor</i>	<b>John Atkinson</b>

---

---

## Editorial Board

---

Richard L. Abel, Professor of Law, University of California, Los Angeles  
Ian Ayres, Professor of Law, Stanford University  
Donald Black, Professor of Sociology, University of Virginia  
Erhard Blankenburg, Professor of Sociology of Law, Free University of Amsterdam  
John Braithwaite, Professor of Law and Sociology, The Australian National University  
Jane F. Collier, Professor of Anthropology, Stanford University  
Phoebe C. Ellsworth, Professor of Law and Psychology, University of Michigan  
Eliot L. Freidson, Professor of Sociology, New York University  
Lawrence M. Friedman, Professor of Law, Stanford University  
Hazel Genn, Professor of Law, Queen Mary and Westfield College, University of London  
Carol J. Greenhouse, Professor of Anthropology, Indiana University  
Valerie P. Hans, Professor of Criminal Justice, University of Delaware  
Earl Johnson, Jr., Associate Justice, California State Court of Appeal  
Doreen McBarnet, Senior Research Fellow, Centre for Socio-Legal Studies, Wolfson College, Oxford  
Carrie J. Menkel-Meadow, Professor of Law, University of California, Los Angeles  
Laura Nader, Professor of Anthropology, University of California-Berkeley  
Judith Resnik, Professor of Law, University of Southern California  
Daniel Rubinfeld, Professor of Economics, University of California-Berkeley  
Dietrich Rueschemeyer, Professor of Sociology, Brown University  
Martin M. Shapiro, Professor of Law, University of California-Berkeley  
Susan S. Silbey, Professor of Sociology, Wellesley College  
David Sugarman, School of Law, Lancaster University  
David M. Trubek, Professor of Law, University of Wisconsin  
Barbara Yngvesson, Professor of Anthropology, Hampshire College  
Michael Zander, Professor of Law, Department of Law, London School of Economics

---

## A Libel Story: *Sullivan* Then and Now

Elena Kagan

ANTHONY LEWIS, *Make No Law: The Sullivan Case and the First Amendment*.  
New York: Random House, 1991. Pp. 354. \$ 25.00.

*New York Times v. Sullivan*<sup>1</sup> is one of those rare cases—perhaps especially rare in the field of First Amendment law—in which the heroes are heroes, the villains are villains, and everyone can be characterized as one or the other. In *Sullivan*, lofty principle need not wrestle with distasteful fact; the ideal of free speech need not come to terms with base or injurious utterance. There is a beautiful simplicity about the case—a stark clarity—that lends itself to a certain brand of storytelling.

Anthony Lewis, columnist and former Supreme Court reporter for the *New York Times*, ranks by any measure among the premier legal storytellers of our time. Almost three decades ago, his *Gideon's Trumpet* made a folk hero of Earl Gideon and turned *Gideon v. Wainwright* into a metaphor for the wise and just use of law. Now Lewis has focused his sights closer to home, telling the equally significant story of how his newspaper in the *Sullivan* case helped to transform the law of libel and the very meaning of the First Amendment.

Among Lewis's talents is the journalistic gift of knowing a good story when he sees one. *Sullivan*, like *Gideon*, has a power stemming from its simplicity. And again like *Gideon*, it oozes drama. For purposes of narrative, it is hard to better a case that involves the most glamorous part of the

---

Elena Kagan is assistant professor of law at the University of Chicago.

The author wishes to thank Al Alschuler, Mary Becker, Anne-Marie Burley, Richard Epstein, Stephen Gilles, Abner Greene, Larry Lessig, Michael McConnell, Martha Minow, Geoffrey Stone, Cass Sunstein, and (especially) David Strauss for valuable comments on earlier drafts and the Bernard G. Sang Faculty Fund and the Bernard Meltzer Fund for financial assistance.

1. 376 U.S. 254 (1964).

© 1993 American Bar Foundation.  
0897-6546/93/1801-0197\$01.00

197

CLINTON LIBRARY PHOTOCOPY

Constitution, arose from the crucible of the civil rights struggle, and produced a decision of historic proportions. But the storyteller here is as good as the story. Pitching his book to a wide audience, Lewis makes not only *Sullivan* but the whole of First Amendment law come alive in the senses and imagination. Lewis explains legal concepts to the general reader with great facility. And he writes from the heart, communicating material that means much to him with the power and emotion necessary to render the material meaningful to others.

*Make No Law* is in part simple narration, but its recitation of facts is virtue, not vice. In telling the story of *Sullivan*, Lewis performs the signal task of demonstrating how much facts matter, of showing the extent to which a legal decision may (and should) be dependent on context and circumstance. Lewis's narrative, including his account of the Supreme Court's deliberations (an account based largely on Justice Brennan's private papers), serves to highlight the role of anecdote in law. In deciding *Sullivan*, the Court was in large part responding to a story—the selfsame story Lewis tells with a journalist's eye for vignette and detail.

Lewis, however, is not content to give just the facts; he spins stories with morals. Juxtaposed against Lewis's immersion in context is a tendency to generalize broadly from his subject matter. On one page, Lewis revels in the particular facts of *Sullivan*; on another, he uses these facts as springboard to justify principles of libel law and First Amendment law applicable to a much wider range of cases. This method, of course, is not itself mistaken. Stories often teach general lessons, and the drawing of morals may be especially appropriate in legal stories because the technique mirrors the way law naturally (perhaps inevitably) develops. Indeed, the Supreme Court has done exactly what Lewis favors, extending the rules articulated in *Sullivan* to a broader set of libel cases and using *Sullivan* as authority for some sweeping First Amendment principles.

The real question is not whether to go beyond the original context of a case but how best to do so. The drawing of morals from a story may be more or less apt; so, too, may be the creation of legal rule and principle. Thus, the question, put more precisely, is whether Lewis—or, more important, the Court—has generalized appropriately from *Sullivan*, has seen what the case was truly about and has used this understanding to denote where the case has relevance.

The answer to this question, I think, is mixed, and in much of the rest of this review I offer some thoughts about how and why this is so. After reviewing Lewis's account of *Sullivan*, I discuss aspects of the decision that raise greater problems than Lewis concedes. I then address two different levels of generality on which the *Sullivan* decision may operate. On the first level, *Sullivan* generates special rules of defamation law; on the second level, discussed more briefly, *Sullivan* stands for broader First Amendment

principles. The use of *Sullivan*, by Lewis and the Court, to support a corpus of defamation law strikes me as troubling: as an effort to fit the square pegs of many defamation cases into the round holes of *Sullivan*. Lewis and the Court, I think, do far better when invoking *Sullivan* on the broader level of First Amendment principle. It is there that the *Sullivan* case resounds most deeply.

## I. THE CASE AND THE RULING

To evaluate when and where the *Sullivan* decision has meaning, the place to start is with the case decided. Lewis provides a vivid account of the underlying controversy as well as its treatment in the courts. This account provides a basis for reflecting on the central concerns of the decision. But what seems to emerge is something different from what Lewis (or any other proponent of expanded protection for libel defendants) might have intended. The context of the case and the Supreme Court's own deliberations suggest that *Sullivan* was only secondarily—almost accidentally—a decision about the law of defamation. The Court's decision—including the puzzling adoption of the actual malice standard—responded primarily to the core First Amendment problem of the abuse of power to stifle expression on public issues, a problem only contingently related to the law of defamation.

### A. The Case

The basic facts of *Sullivan* are familiar, although perhaps more so to readers of this journal than to readers of Lewis's book. On 29 March 1960 an advertisement sponsored by the Committee to Defend Martin Luther King and the Struggle for Freedom in the South appeared in the *New York Times*. The ad, headlined "Heed Their Rising Voices," contained ten paragraphs of text detailing efforts by "Southern violators of the Constitution," including police officers, to derail the civil rights struggle through acts of governmental abuse and violence. L. B. Sullivan, a Commissioner of the City of Montgomery, Alabama, in charge of supervising the city's police, brought a libel suit based on the ad against the *New York Times*. Sullivan's name never appeared in the ad, but he claimed that statements about the Montgomery police and southern law violators had been read to refer to him. Sullivan further claimed that several admitted—though mostly minor—inaccuracies in the ad had harmed his reputation. An Alabama jury returned a verdict for Sullivan in the full amount demanded—a half-million dollars—and the Alabama Supreme Court affirmed. So much is found in the U.S. Supreme Court's majority opinion.

To this factual core, Lewis adds a wealth of detail about the case and its treatment in the Alabama trial court. The advertisement quickly became notorious in Montgomery, even though only about 400 copies of the *New York Times* were circulated in all of Alabama. The *Montgomery Advertiser*, the morning newspaper in the city, brought the ad to the city's attention with an editorial charging "crude slanders against Montgomery" by "voluntary" and "involuntary liars" (at 11); and by the time the *New York Times* went searching for local counsel, not a single Montgomery lawyer would take the case (at 24). (The *Times* found a Birmingham lawyer, who booked hotel reservations for the primary *Times* counsel, Lewis Loeb, under an assumed name (at 24).) The trial judge, Walter Burgwyn Jones, publicly had proclaimed his belief in "white man's justice" and had authored a tract entitled *The Confederate Creed*; on the 100th anniversary of the founding of the confederacy, he had participated in a reenactment of the swearing in of Jefferson Davis (administering the oath of office) and then retired to his courtroom to preside over a trial in which jurors wore confederate uniforms (at 25-26). Lewis conjectures that Jones may even have helped to plan Sullivan's libel suit, although (as Lewis admits) there is little in the way of proof to back up this surmise (at 27). In any event, Jones presided over the trial (in a racially segregated courtroom) and found in favor of the plaintiff on every significant ruling; the all-white jury, instructed that the advertisement was libelous, false, and injurious as a matter of law, took about two hours to decide that the advertisement was "of and concerning" Sullivan and that he should receive \$500,000 (at 32-33).

Lewis shows that the *Sullivan* trial was merely the first salvo in a concerted campaign against the northern establishment press by southern public officials and opinion makers—a campaign which intended to curtail media coverage of the civil rights struggle and threatened to succeed in this design. The *Sullivan* case was the first of five suits brought by public officials based on the "Heed Our Rising Voices" advertisement; each of the other suits also claimed damages of \$500,000 (at 35). Two stories by *New York Times* reporter Harrison Salisbury prompted another round of libel suits, asking for total damages of \$3,150,000 against the *Times* and \$1,500,000 against Salisbury (at 22). Nor was the *Times* the only target; by the time the Supreme Court decided *Sullivan* in 1964, southern officials had brought nearly \$300 million in libel actions against the press (at 36). The *Montgomery Advertiser* candidly headlined a story about the libel cases: "State Finds Formidable Legal Club to Swing at Out-of-State Press" (at 35). The *Alabama Journal*, Montgomery's evening paper, noted that the suits "could have the effect of causing reckless publishers of the North . . . to make a re-survey of their habit of permitting anything detrimental to the South and its people to appear in their columns" (at 34). And the suits indeed could have had great effect. The *Times* withdrew all of its reporters

from Alabama for a year in order to maintain a personal jurisdiction argument (at 43). There was some danger that the newspaper, then struggling with labor disputes and making minuscule profits, would not survive (at 35). CBS, according to one of its attorneys, would have ceased doing programs on the southern civil rights movement had the *Sullivan* verdict not been reversed (at 245).

Lewis leaves little doubt that the particular facts of *Sullivan*, as well as the surrounding libel suit campaign, powerfully affected the Court. Suppose, Lewis asks the reader to consider, *Sullivan* had had the modesty (or foresight) to strike a zero from his damage claim (at 161). Would the Court then have decided to review the case? Would the *Times* even have filed a cert petition?<sup>2</sup> Or suppose that the *Times* advertisement really had defamed *Sullivan*, referring to him by name in a manner that unjustly harmed his reputation.<sup>3</sup> Or suppose that the *Sullivan* suit had not instigated a flood of other libel cases by southern officials—cases specifically noted in both the majority opinion and Justice Black's concurrence. Lewis, as well as all of the attorneys involved in the case, believe that the Court would have let the *Sullivan* verdict stand in the absence of this special set of circumstances (at 161): what galled the Court was something much more than that a single public official had recovered a libel judgment for an innocent defamatory statement.

## B. The Ruling

One clue to understanding what concerned the Court in *Sullivan* may lie in an arresting quiet at the center of the case—specifically, in the Justices' failure during deliberations to criticize, debate, or question the majority opinion's adoption of the actual malice standard.<sup>4</sup> Although Lewis

---

2. Lewis notes that soon after the *Times*'s general counsel requested Herbert Wechsler to draft a petition for certiorari, *Times* editors summoned Wechsler to a meeting to defend the decision to seek review of the verdict. Wechsler told Lewis: "I was being asked to show cause why I should file a petition for certiorari. I found myself defending the legal position I was advancing in defense of the *Times*—that the First Amendment applied to libel cases. . . . People were asking why it wasn't enough for the *Times* to 'stick to our established position that we never settle libel cases, we publish the truth, if there's an occasional error we lose and that's one of the vicissitudes of life'—that at a time when I was told the paper was barely making a profit and these judgments were mounting up" (at 107). The anecdote reveals how greatly the attitudes and expectations of the American press have changed since *Sullivan*—perhaps due to *Sullivan* itself. See *infra* at pt. II.A. pp. 711–12.

3. Justice Black's concurring opinion in *Sullivan* wryly noted that the "record lends support to an inference that instead of being damaged Commissioner Sullivan's political, social, and financial prestige has likely been enhanced by the *Times*' publication." 376 U.S. at 294.

4. Under the actual malice rule, a libel plaintiff must show that the defendant published the challenged statement "with knowledge that it was false or with reckless disregard of whether it was false or not." *Id.* at 280. A defendant acts with "reckless disregard" when he "in fact entertain[s] serious doubts as to the truth of his publication." *St. Amant v.*

fails to highlight the point, his description of Justice Brennan's papers (which include all the Justices' notes to each other, as well as a memorandum written by a law clerk detailing the Court's deliberations) makes this lack of controversy immediately apparent. Throughout the Court's lengthy and active consideration of the case,<sup>5</sup> the actual malice standard—in hindsight, by far the most significant aspect of the opinion—occasioned almost no debate, even though the Court at conference had agreed to decide the case on the narrower ground that the Constitution required clear and convincing proof of every traditional element of a libel action in a case involving a public official (at 165–66). Justices Douglas, Black, and Goldberg sent notes to Justice Brennan explaining that they favored absolute immunity from libel suits brought by public officials about official conduct (at 171), and they eventually filed concurring opinions taking this position. But none of the Justices who ultimately signed on to Justice Brennan's opinion raised any questions about the actual malice standard: Was it too strict? Was it necessary? Where did it come from? How would it work? On these critical issues, silence reigned.

The Justices instead fretted about a part of the opinion that today seems far less important: the application of the new standard to the evidence in the case. All the notes from the Justices concerned the question whether the Court properly could apply the actual malice standard to the evidence below—or could go even further and prevent a new trial at which Sullivan might offer additional evidence (at 172–82).<sup>6</sup> The latter position attracted little support, but the former eventually trumped concerns that the Court would overreach its authority by examining the sufficiency of the evidence under the new standard. Chief Justice Warren noted that without such an examination, “we will merely be going through a meaningless exercise. The case would be remanded [and] another improvisation would be devised” (at 178). And Justice Brennan reminded his colleagues that “there are a number of other libel suits pending in Montgomery and in Birmingham and those concerned should know what to expect in the way of judicial superintendence from this Court” (at 177). Such pragmatic concerns about applying the new rule to prevent recovery by Sullivan and others—about dealing with the particular problem facing the Court—overwhelmed abstract consideration of the rule itself.

But more might be said than that the factual situation before the Court pushed legal questions to the margin: the adoption of the actual

---

Thompson, 390 U.S. 727, 731 (1968). Under traditional common law rules of libel, the plaintiff need make no such showing: the plaintiff must establish only that the defendant has published a false defamatory statement “of and concerning” the plaintiff.

5. Justice Brennan wrote no fewer than eight drafts of his majority opinion, most of which were circulated to the other members of the Court (at 164).

6. The alternative, of course, was to leave to the state courts the task of applying the new standard to the evidence in the case.

malice rule by Justice Brennan, and the Court's ready and unquestioning acceptance of it, may in fact have resulted from the extraordinary circumstances of the case. One of the great puzzles of *Sullivan* concerns why the Court adopted the actual malice rule rather than decide the case on one of numerous available grounds based on common law principles: that the published statements were not "of and concerning" Sullivan; that they were not substantially false; that they did not injure his reputation. Richard Epstein, for example, recently has suggested that the *Sullivan* Court took the wrong tack: that it should have decided the case on the ground that the common law rules of libel represent the constitutional norm in a public official libel case and that the Alabama courts had failed to follow these rules.<sup>7</sup> Justice Brennan's initial rationale for reversing the judgment—that the Constitution requires clear and convincing proof of every traditional element of defamation in actions involving public officials—resembles Epstein's proposed approach: each imports the Constitution into libel cases brought by public officials, but in some manner pegs constitutional requirements to the common law. *Make No Law* does not reveal precisely why Justice Brennan abandoned his initial rationale and adopted the actual malice rule: Lewis says only that Brennan wrote the initial draft himself and must have changed his mind in the course of composition (at 166). The broader story that Lewis tells, however, may provide the key—and it may do so in either of two related ways.

Most pragmatically, if the dominant concern of the Court was to prevent recovery not only by Sullivan but by the host of other southern officials who had filed libel suits on the basis of articles about the civil rights movement, the actual malice standard may have appeared by far the best approach. Even if the *Sullivan* verdict itself could have been reversed by constitutionalizing common law rules, numerous other libel cases brought by southern officials—some undoubtedly stronger under common law principles—would remain. The Court's decision in *Sullivan* removed the threat of all these cases: Sullivan himself decided not to seek a new trial and the other libel actions brought by southern officials quickly fell away (at 161). It seems doubtful whether Justice Brennan's original rationale or any other similar approach would have sent so strong a message or had such a powerful effect. Without some significant addition to common law requirements, the Court may have felt, the danger confronting speech about the civil rights movement would not dissipate. As Justice Black wrote in a note to Justice Brennan lauding his opinion, "Most inventions even of legal principles come out of urgent needs" (at 175).

On a somewhat deeper level, the adoption of the actual malice standard may have resulted from the Court's understanding that the "ur-

---

7. See Richard Epstein, "Was *New York Times v. Sullivan* Wrong?" 53 *U. Chi. L. Rev.* 782, 792-93 (1986).

of failure to prove actual malice. He thus became one of the first victims of the *Sullivan* standard: persons who (unlike Sullivan himself) had suffered real reputational injury and yet were unable to recover for it.

Lewis's response to these cases is the familiar one (implicitly adopted in *Sullivan*) that such personal harm is the unfortunate but necessary consequence of a rule promoting the social good of uninhibited comment concerning public officials. But even if we assume that the actual malice standard in fact encourages speech about public officials—itsself a somewhat uncertain proposition<sup>13</sup>—this response begs an important (if almost equally familiar) question: Is uninhibited defamatory comment an unambiguous social good? That is, does it truly enhance public discourse?<sup>14</sup>

This question, never addressed by Lewis, poses a challenge to his and the *Sullivan* Court's view of the effect of the actual malice standard, outside the context of the *Sullivan* case itself, on the quality of public discourse and hence on the democratic process. The ultimate concern of *Sullivan* was to strengthen that process by ensuring that the citizenry receive important information about the conduct and policies of government officials. Certainly, the application of the actual malice standard in *Sullivan* served that function. But the malice standard may not have the same effects when applied more generally. Several commentators have noted that to the extent *Sullivan* decreases the threat of libel litigation, it promotes not only true but also false statements of fact—statements that may themselves distort public debate.<sup>15</sup> Here, too, the *Goldmark* case provides a telling counterexample to *Sullivan*: the false charges of Communist Party associations in that case more likely corrupted than enhanced the realm of public

---

13. Lewis devotes much attention to whether the actual malice rule actually encourages speech, but his discussion backtracks on itself. On the one hand, Lewis insists that the *Sullivan* rule was responsible for press coverage of some of the most important national stories of the past decades, including Watergate and the Vietnam War (at 158). The part of this claim relating to Vietnam seems wildly overdrawn. If, as Lewis writes, journalists during the Vietnam War began to show less deference to official accounts and judgments than in the earlier years of the Cold War, surely this newfound independence had more to do with changed attitudes toward government players and policies than with changed rules of libel law. Lewis seems on more solid ground when he contends that libel rules affected coverage of the Watergate scandal. Still greater plausibility would attach to a claim that the actual malice rule freed smaller media outlets, whose very existence could be threatened by a libel judgment, to confront powerful local politicians. But Lewis makes a number of observations that place even this scaled-down claim in doubt. He notes that several earlier periods of American history saw savage attacks on political leaders by the press (at 206–7); and he concludes that the “notion that the press was harder on public servants after 1964 is contradicted by history” (at 206). Even more important, Lewis several times asserts (in making the claim for augmented libel protection) what has become a commonplace in press circles: that in practice the actual malice rule has raised the costs and stakes of libel litigation and thereby may have increased press inhibitions (at 200–202, 244). In the absence of any empirical data, choosing between such rival assertions becomes a matter of crude intuition.

14. The question is discussed in most expansive form in Lee Bollinger, *Images of a Free Press* 26–39 (Chicago: University of Chicago Press, 1991).

15. See, e.g., *id.* at 26–27, 35–36.

discourse. In this way, the legal standard adopted in *Sullivan* may cut against the very values underlying the decision.

The problem, indeed, may go even deeper: it may involve not merely the promotion of false statements but also a more general tendency to sensationalize political discourse. When the press stops worrying about the accuracy of defamatory statements, it may start covering subject matter not readily amenable to determinations of truth or falsity; that subject matter, whether true or false, often ranks high in sensationalist content. Thus, the *Sullivan* decision, although itself involving core political speech, may have facilitated (which is not to say "caused") both the rise of tabloids and the "tabloidization" of the mainstream press. And arguably, such expression—the obvious example here is speech concerning the private and sexual lives of political figures—distracts from and devalues the kind of discourse *Sullivan* meant to promote. The poverty of such speech does not itself provide a reason for suppression; the First Amendment would mean little if government could restrict speech whenever it were deemed distracting or demeaning or even false. But with respect to libel law, the interest in reputation provides the reason for regulation; the regulation falls only because the benefits of the additional speech outweigh its reputational costs. To the extent that the speech promoted makes little contribution to public dialogue, the relaxation of libel law seems difficult to countenance.

*Make No Law* includes copious evidence that the press in pre-*Sullivan* days demonstrated great sensitivity to this range of questions. Lewis, for example, recounts that just after the Court decided *Sullivan*, a principal editor of the *Times* wrote a letter to Herbert Wechsler, author of the *Times*'s winning brief, saying that "we may be opening the way to complete irresponsibility in journalism" and asking whether it was right to erode principles of journalistic responsibility just "because justice is lopsided in one area of the nation" (at 219–20). Similarly, Lewis notes the reluctance of the *Times* even to ask the Supreme Court to review the *Sullivan* case given the newspaper's standard position that "we publish the truth, if there's an occasional error we lose and that's one of the vicissitudes of life" (at 107). And Lewis cites several cases in which the press approved of libel verdicts notwithstanding (or even because of) their inhibition of speech: in one case, the *Times* praised a \$3.5 million judgment as likely to have a "healthy effect" on public discourse (at 112).

Today's press engages in far less examination of journalistic standards and their relation to legal rules. Rather than asking whether some kinds of accountability may in the long term benefit journalism, the press reflexively asserts constitutional insulation from any and all norms of conduct. Lewis himself notes this air of exceptionalism and entitlement. He writes that "[p]hrases such as 'freedom of the press' or 'First Amendment rights'" have assumed the aspect of "exclusivist dogma[,]" with "[s]ome

editors and publishers act[ing] as if the . . . First Amendment were designed to protect journalism alone, and to make that protection superior to other rights" (at 208). In this vein, he aptly observes: "When the Supreme Court decides a case against a claimed press interest, editors and publishers too often act as if the Constitution were gone" (at 209). And Lewis discusses as well the unwillingness of the press to confess to error, using as an example of this "stiff-necked press behavior" *Time* magazine's refusal to retract an unsupported assertion about the activities of former Israeli Defense Minister Ariel Sharon (at 208). Yet having said all this, even Lewis stops short of questioning whether current libel law has had any detrimental effects on journalistic practice. What *kinds* of speech has *Sullivan* promoted? Is it related—and, if so, how—to trends in journalism of dubious value? Lewis, unlike the earlier generation of journalists he cites, declines to consider these old (but never more potent) issues.

And this contrast raises a final question about the unintended effects of *Sullivan*: Is it possible that *Sullivan* bears some responsibility for a change in the way the press views itself and its conduct—a change that the general public might describe as increased press arrogance? It is wise to be wary about attributing too much cultural impact to a Supreme Court decision; yet it is hard to believe that those most directly affected by a decision like *Sullivan* are in no way changed by it. At the most basic level, judicial declarations of unaccountability can go to the head. It is hardly unthinkable that increased legal protection may lead to a greater sense of entitlement and self-importance (which in turn may manifest itself in questionable conduct). But the effects of *Sullivan* on the press's conception of itself may go yet deeper. Just as the Court treats the story of *Sullivan* as an archetype, so too may the press: the heroic role of the *Times* in that case helps to define and inform self-understanding. This mythical image may at times serve as model, but it also may blind the press to numerous less attractive aspects of its role and performance.<sup>16</sup> Thus, the self-image of the press becomes semi-delusional, and journalists cease to ask the questions of themselves which they ask of other powerful actors in society.

Questions of this kind in no way prove that the Court decided *Sullivan* incorrectly or that the Court now should reconsider its holding. The story of *Sullivan* rebels against this conclusion, whether that story is framed as a particular tale of how southern public officials attempted to suppress commentary about the civil rights movement (and thus to suppress the

---

16. In *Images of a Free Press*, Lee Bollinger posits that the image of the press portrayed in *Sullivan* and similar cases may entice the press to conform to norms of quality journalism. See *id.* at 40–61. I agree with Bollinger that the *Sullivan* Court articulated a certain image of the press and that the press largely has absorbed that image. We disagree as to the consequences of this process. Whereas Bollinger believes that the absorption of the *Sullivan* image often uplifts the press, I believe the absorption of that image more often succeeds only in blinding the press to its own shortcomings as well as its capacity to inflict unjust harm.

movement itself) or as a more general tale of how government officials may attempt to stifle criticism of themselves and their policies. But to view *Sullivan* as a kind of icon—a decision about which “nothing more need be said” (at x)—is too easy by half. If nothing else, such a view may distort consideration of the question whether and how *Sullivan* should be extended. This question has occupied the Court from the time of *Sullivan* to this day, and Lewis discusses the Court’s responses in detail. But because he fails to acknowledge fully the difficulties associated with the *Sullivan* rule itself, he can accept in the blandest way all further extensions of the principle. He need never confront the question—a question intertwined with the very meaning of *Sullivan*—of the decision’s proper limits.

## B. Questionable Extensions

In one of the first commentaries on *Sullivan*, Harry Kalven predicted that the Court would not long view the decision as “covering simply one pocket of cases, those dealing with libel of public officials.”<sup>17</sup> The Court, Kalven predicted, would accept an “overwhelming . . . invitation to follow a dialectic progression” from the category of public officials to other categories yet further-reaching.<sup>18</sup> And although Kalven mistook the precise steps in the progression,<sup>19</sup> he soon saw confirmed his basic prophecy. In a series of cases succeeding *Sullivan*, the Court extended at least some level of constitutional protection to defendants in nearly all libel cases. This course, however, proved more problematic than Kalven anticipated. In extending *Sullivan*, the Court increasingly lost contact with the case’s premises and principles. Even when viewed most broadly, *Sullivan* relied upon two essential predicates: a certain kind of speech and a certain kind of power relationship between the speaker and the speech’s target. These attributes of the case, once so vital, became submerged in the Court’s subsequent construction of libel doctrine.

The constitutional scheme that today governs libel cases is familiar, the way it operates in practice somewhat less so. Not only public officials, but also so-called public figures must prove actual malice to recover for defamation.<sup>20</sup> Who is a public figure? Although the Supreme Court attempted for some years to impose limits on the category, lower courts have interpreted it expansively.<sup>21</sup> The official definition of a public figure in-

17. Harry Kalven, Jr., “The *New York Times* Case: A Note on ‘The Central Meaning of the First Amendment,’” 1964 S. Ct. Rev. 191, 221.

18. *Id.*

19. Kalven believed that in offering constitutional protection from libel suits, the Court should and would move “from public official to government policy to public policy to matters in the public domain.” *Id.*

20. *Curtis Publishing Co. v. Butts*, 388 U.S. 130 (1967).

21. The Supreme Court has ruled on several occasions that libel plaintiffs were not

cludes any individual who has achieved "pervasive fame or notoriety" or who "voluntarily injects himself or is drawn into a particular public controversy."<sup>22</sup> A more informal definition might go something like: Everyone the reader has heard of before and a great many people he hasn't. The vast majority of those likely to attract media attention fall within the category.<sup>23</sup> Even plaintiffs lucky enough to be labeled private figures generally must satisfy a heightened standard of liability: negligence to recover compensatory damages and actual malice to obtain presumed or punitive damages.<sup>24</sup> Given the frequent difficulty of proving actual injury in defamation suits, as well as the economics of litigation, many of these suits are tenable only with evidence of actual malice.<sup>25</sup> In a tiny category of cases, in which a private figure is defamed on a "matter of purely private concern," the actual malice standard disappears, as may all other constitutional requirements.<sup>26</sup> The upshot of the system is that the constitutional standard established in *Sullivan* for a public official bringing a libel suit against critics of his official conduct today governs the bulk of defamation cases, at least against media defendants.<sup>27</sup>

Lewis is generally sanguine toward this result, though he admits now and again to some concern. Tracing the course of post-*Sullivan* libel law, Lewis arrives at the case of entertainer Wayne Newton, who lost a multi-million-dollar libel judgment, arising from an allegation that he associated with a Mafia figure, for failure to prove actual malice (at 197-98). "Philosophically," Lewis concedes, "cases like Wayne Newton's are a long way from the Alabama lawsuit that led the Supreme Court to bring libel within the First Amendment"; had the *Sullivan* case never arisen, the Supreme

---

public figures. See *Wolston v. Reader's Digest*, 443 U.S. 157 (1979); *Hutchinson v. Proxmire*, 443 U.S. 111 (1979); *Time, Inc. v. Firestone*, 424 U.S. 448 (1976); *Gertz v. Robert Welch, Inc.*, 418 U.S. 323 (1974). For a review of lower court decisions to the opposite effect, see David Anderson, "Is Libel Law Worth Reforming?" 140 *U. Pa. L. Rev.* 487, 500-501 (1991).

22. *Gertz v. Robert Welch, Inc.*, 418 U.S. at 351.

23. See Anderson, 140 *U. Pa. L. Rev.* at 501. Some cases demonstrating the range of the public figure category are: *Trotter v. Jack Anderson Enterprises*, 818 F.2d 431 (5th Cir. 1987) (president of Guatemalan soft-drink bottling company); *McBride v. Merrell Dow & Pharmaceuticals*, 800 F.2d 1208 (D.C. Cir. 1986) (expert witness); *Dameron v. Washington Magazine*, 779 F.2d 736 (D.C. Cir. 1985) (air traffic controller); *Brewer v. Memphis Publishing Co.*, 626 F.2d 1238 (5th Cir. 1980) (former girlfriend of Elvis Presley), *cert. denied*, 452 U.S. 962 (1981); *James v. Gannett Co.*, 353 N.E.2d 834 (N.Y. 1976) (belly dancer).

24. *Gertz v. Robert Welch, Inc.*, 418 U.S. at 347-49.

25. See Anderson, 140 *U. Pa. L. Rev.* at 502.

26. See *Dun & Bradstreet v. Greenmoss Builders*, 472 U.S. 749, 759-61 (1985), in which the Court held that in a private figure/private concern case, a showing of actual malice was unnecessary even to obtain presumed and/or punitive damages. The Court left open the question whether any heightened constitutional standards (relating, for example, to burdens of proof) apply in such cases.

27. One study of appellate cases involving media defendants found 75 cases in which the actual malice standard controlled and only 24 in which any lesser standard controlled. See Marc Franklin, "Suing Media for Libel: A Litigation Study," 1981 *A.B.F. Res. J.* 795, 824.

Court surely would have rebuffed a claim by the press that the First Amendment barred Newton's recovery (at 198). But having acknowledged the gulf between the two cases, Lewis minimizes its significance. What took the Court from protecting the "citizen-critic of government" to protecting detractors of an entertainer, Lewis writes, was good constitutional common law decision making. In expanding the reach of the First Amendment over libel actions, the Justices "were guided by their sense of the society: its traditions, its needs, its changing character" (at 198). Indeed, Lewis implies, we should count ourselves fortunate that "a libel case that really did engage the central meaning of the First Amendment had come along" first, so that it could call forth "a transforming Supreme Court decision" which then would "spread to a much larger field" (at 198).

But the mere restatement of this conclusion demonstrates its oddity. Why is the Supreme Court's libel jurisprudence an example of praiseworthy incremental decision making if it extended constitutional protection from cases that "really did engage the central meaning of the First Amendment" to cases that (impliedly) really did not? What does a case concerning criticism of a government official's public conduct have to do with a case concerning comments on a popular entertainer's private associations? The chasm between the two cases noted by Lewis easily could have been even wider. The actual malice standard would have applied in *any* libel suit brought by Newton, simply by virtue of his fame;<sup>28</sup> imagine, for example, a case arising from an allegation not of a Mafia connection but of an adulterous relationship. Or consider the many cases—the recent *Masson v. New Yorker* is an example—in which the actual malice standard applies even though the plaintiff is both unknown beyond a narrow circle and uninvolved in governmental affairs, because of his participation in one of the countless significant and not-so-significant matters that can be deemed a "public controversy."<sup>29</sup> The use of the actual malice standard in this wide range of cases appears to have little connection with the story of *Sullivan*. Viewed from that vantage point, current libel law seems the result not of steady and sensible common law reasoning but of a striking disregard of the doctrine's underpinnings.

28. "Pervasive fame or notoriety" makes a person a public figure for purposes of any statement made about him, regardless of the subject matter. See *Gertz v. Robert Welch, Inc.*, 418 U.S. at 351. Newton, like all "celebrities," thus qualifies as a public figure in any libel suit. Indeed, the district court handling Newton's case imposed sanctions of \$55,000 on him for contesting the public-figure issue at all. See *Newton v. National Broadcasting Co.*, 930 F.2d 662, 668 n.6 (9th Cir. 1990).

29. See *Masson v. New Yorker Magazine*, 111 S. Ct. 2419 (1991). Although in one case the Supreme Court held that a widely publicized divorce proceeding was not a public controversy, see *Time, Inc. v. Firestone*, 424 U.S. 448 (1976), courts generally have shied away from declaring matters reported on by the press to be something other than public controversies. The more usual ground for private-figure status is that the plaintiff insufficiently involved himself in the relevant controversy. See, e.g., *Gertz v. Robert Welch, Inc.*, 418 U.S. 323; *Wolston v. Reader's Digest*, 443 U.S. 157.

However unwittingly, *Make No Law* thus supports a claim that the Court accorded *Sullivan* too great a significance—a significance outstripping the case's real meaning—in the development of constitutional rules relating to libel law. In the course of constructing a constitutional regime to govern libel actions, the Court assimilated to *Sullivan* an array of cases divergent in character. Factual situations posing different concerns, implicating different principles, became as one. In this sense, the development of libel law may be viewed not as a "rich . . . illustrati[on]" of the common law method in First Amendment law (as Lewis would have it (at 183)), but as a deviation from that method, with its characteristic focus on particulars and their relation to established principles.<sup>30</sup>

To say this much, of course, is not to claim that the Court should have declined to extend *Sullivan* at all. If *Sullivan* is not prototypical of libel actions, neither is it likely to be wholly freestanding. *Sullivan* may well have relevance beyond its boundaries, because libel of government officials may share sufficiently important traits with other instances of libel to justify extension of the actual malice rule to the latter. The key is to identify and explain the relevance of those common attributes. In this regard, two complementary possibilities present themselves.

One approach, articulated in various ways by both the Justices and commentators, would apply the actual malice rule to all (but only) those cases involving speech on governmental affairs—or, stated more broadly, speech on matters of public importance—or, stated still more broadly, speech on matters of public concern or interest.<sup>31</sup> This approach emerges from viewing *Sullivan* as primarily a case about the speech necessary for democratic governance. Such a view draws on some of the most notable features of the *Sullivan* opinion—the emphasis on seditious libel, the concern that citizens have access to the information necessary to act in their intended sovereign capacity, the statement of "profound national commitment to the principle that debate on public issues should be uninhibited, robust, and wide-open."<sup>32</sup> Under this approach, a court would consider whether the speech in question is of a kind similar to the speech in *Sullivan*, in the sense that the content of the speech affects or relates to self-

30. For an exploration of this method, focusing on its use in First Amendment cases, see Cass Sunstein, "On Analogical Reasoning," 106 *Harv. L. Rev.* 741 (1993).

31. See Cass Sunstein, "Free Speech Now," 59 *U. Chi. L. Rev.* 255, 311 (1992) ("The test for special protection should be whether the matter bears on democratic governance, not whether the plaintiff is famous"); Frederick Schauer, "Public Figures," 25 *Wm. & Mary L. Rev.* 905 (1984). Justice Brennan appeared to advocate a similar approach when he urged that the actual malice standard apply to all cases involving speech on "matters of public interest." *Rosenbloom v. Metromedia*, 403 U.S. 29, 42 (1971) (Brennan, J.). But for Justice Brennan, this protection may have been meant to enhance, rather than to replace, the protection automatically accorded in public figure cases.

32. 376 U.S. at 270.

government. Comments, for example, about Wayne Newton or other celebrities currently classified as public figures might well flunk this test.

A second approach might focus not (or not only) on the content of the speech but on a concern arguably as essential to the *Sullivan* decision: the respective power of the speaker and the subject and the relation between the two. In *Sullivan*, the press (to the extent it targeted any particular individuals) criticized persons of substantial influence. Those persons derived their power from government positions, but this fact alone may not be of paramount importance. Chief Justice Warren understood *Sullivan* as resting on the simple presence of power—whether governmental or private did not matter—and the fear of its abuse.<sup>33</sup> To him, the principle of *Sullivan* applied with equal clarity to important figures in the “intellectual, governmental, and business worlds,” both because individuals in each of these spheres exerted influence over the ordering of society and because they alike had means to counter criticism.<sup>34</sup> The implicit comparison is to cases in which speakers—who themselves may possess enormous influence—target individuals of lesser power and prominence. In such cases, the press may appear in the position of the southern officials of *Sullivan*, the targeted individuals in the position of the *New York Times*.

Under this view, the relevant spectrum in libel law runs between a case like *Sullivan* and a case in which the institutional press defames a relatively powerless individual (regardless whether the person might be viewed as involved in a public controversy). In *Sullivan* itself, the *New York Times* had little circulation and less influence in the relevant community; the ostensible target of its speech, by contrast, controlled vital levers of patronage and power. This situation has little in common with such recent Supreme Court cases as *Masson* or *Milkovich*, the former involving a renowned national magazine which allegedly defamed (by misquoting) the former Projects Director of the Sigmund Freud Archives, the latter involving a local newspaper that accused a high school coach of perjury.<sup>35</sup> Nor does *Sullivan* resemble, with respect to considerations of power, a host of cases that never reach the Supreme Court: for example, *Dameron v. Washington Magazine*, in which a magazine charged an air traffic controller with responsibility for a major accident (and subsequently retracted the state-

33. See *Curtis Publishing Co. v. Butts*, 388 U.S. 130, 163–64 (1967) (Warren, C.J., concurring).

34. *Id.* at 163.

35. In *Masson v. New Yorker Magazine*, 111 S. Ct. 2419 (1991), the plaintiff conceded public-figure status at the beginning of the case; the Supreme Court granted certiorari to consider the question whether and when deliberate misquotation could constitute evidence of actual malice. In *Milkovich v. Lorain Journal Co.*, 110 S. Ct. 2695 (1990), the plaintiff initially was held to be a public figure; only on his second appeal to the Ohio Supreme Court, occurring almost a decade after he brought suit, was this determination reversed. The Supreme Court reviewed the case to determine whether libel defendants are constitutionally entitled to a privilege for statements of opinion.

ment);<sup>36</sup> or *Fitzgerald v. Penthouse International*, in which a magazine accused an expert in "dolphin technology" with having committed espionage.<sup>37</sup> Part of what seems troubling about applying the actual malice rule to these cases arises not so much from the content of the speech (whether it relates to democratic governance) as from the respective societal positions of the speaker and the target. In such cases, the law insulates powerful institutional actors—possessing both a great capacity to harm individuals and a far-reaching influence over society at large—from charges of irresponsibility made by persons with little societal influence and few avenues of self-protection. If part of the point of *Sullivan* was to check the abuse of power and to ensure the accountability of those wielding it, then these cases suggest that the Court's constitutionalization of libel law has gone askew.

Assuming this is so, why has it occurred? The harshest interpretation is that the Court too little probed the foundations of *Sullivan* for clues to its proper application. Under this view, the very rightness of the *Sullivan* result combined with the power of its rhetoric to distract the Court from the (once all-important) context of the decision. The mismatch between *Sullivan* and many current libel cases is due simply to a lack of care and attention in applying the decision.

But a deeper explanation is available, involving the perceived necessity of using categorical rules in libel cases. For reasons having to do with certainty and predictability, the Court often has abjured contextual case-by-case inquiry in First Amendment adjudication, preferring to create rules applicable to broad categories of cases. Once a determination is made to adopt this approach in libel law, the question how to define the categories presents itself. Factors like those I have considered—the connection of speech to self-government or the relationship between the power of a speaker and a subject—resist reduction to simple categorical rules. Even if we were sure that power relations were all that mattered, how could we frame a rule to capture and compartmentalize so elusive a thing as the "power" of a speaker or a subject, let alone the relationship between the two? How could we then incorporate into this rule consideration of the content of the speech and its relation to democratic government—an inquiry which itself appears to demand a kind of fine discrimination in tension with the technique of categorization? However a flat rule is articulated, it may seem inadequate to the task to be accomplished.

The failures of the Court's libel law decisions ultimately may derive from just these problems rather than from a simple failure to respect the underpinnings of *Sullivan*. In some sense, the Court's categorical rules reflect an understanding of *Sullivan* as a case concerning both self-government and power relations. The public figure/private figure dichotomy is

---

36. 779 F.2d 736 (D.C. Cir. 1985).

37. 691 F.2d 666 (4th Cir. 1982).

animated partly by power considerations;<sup>38</sup> and the very definition of a public figure, as well as the bifurcation of the private figure sphere, reflect the view that some kinds of speech are of greater public importance than others.<sup>39</sup> But of necessity a great deal has been lost in the Court's attempt to combine and conflate these highly contextual considerations into a single set of categorical rules, susceptible of ready and predictable application. Because the rules serve only as rough and incomplete proxies for in-depth analysis of the factors relevant to *Sullivan*, the results as often as not fail to comport with the origins of libel law doctrine. In short, what has been lost in the Court's creation of our current highly stylized libel law regime—although perhaps inevitably—is *Sullivan* itself.

### III. SULLIVAN AND FIRST AMENDMENT PRINCIPLE

And yet, on a different level, *Sullivan* may be counted (as I think Lewis would count it) the Court's most successful First Amendment decision. *Sullivan* may have proved a problematic foundation for libel law; it may differ too greatly from most (or many) libel cases to provide a sensible doctrinal base. But the very facts that make *Sullivan* an oddity in libel law place it in the mainstream of First Amendment law generally. As Justice Brennan may have recognized in writing the *Sullivan* opinion, the facts of *Sullivan* present in dramatic form the central concern of the First Amendment: the use of power—most notably, though not exclusively, government power—to stifle speech on matters of public import. Thus, *Sullivan* has served as an utterly reliable source not of libel doctrine but of broad First Amendment principle. And it is in making this point, in operating at this highest level of generality, that *Make No Law* truly shines.

The strongest portions of *Make No Law*, aside from the narration of the *Sullivan* story itself, come when Lewis leaves the field of libel law behind him and focuses on the broader wellsprings and offshoots of the decision. Lewis performs the prodigious feat of describing in an accessible but never simplistic way the development of the major principles of First

38. The Court has provided two justifications for treating public figures differently from private figures, one involving the greater self-help remedies available to public figures and the other involving public figures' assumption of risk. See *Gertz v. Robert Welch, Inc.*, 418 U.S. at 344. At least the former justification is related to the power relations concern explicated in the text. Only Chief Justice Warren, however, has discussed explicitly the question of power in libel cases. See *Curtis Publishing Co. v. Butts*, 388 U.S. at 163-64.

39. Determination of public figure status partly involves the question whether the defamation has arisen from the plaintiff's participation in a "public" controversy, a term which at least suggests an inquiry into the subject matter of the speech. See *supra* text at note 22. Moreover, the private-figure sphere is itself divided into two subcategories by reference to whether the speech concerns public or "purely private" matters. See *supra* text at notes 24-26.

Amendment jurisprudence, in which *Sullivan* played a role central in both chronology and importance. If at times the narration smacks of Whig history, as the Court progresses ever onward toward the most luminous of goals, the material may provide ample justification for the treatment. Thus, Lewis links *Sullivan* on the one end with the great dissenting and concurring opinions of Holmes and Brandeis,<sup>40</sup> as well as with *Near*,<sup>41</sup> *Grosjean*,<sup>42</sup> and *Bridges*,<sup>43</sup> and on the other end with *Bond*,<sup>44</sup> *Brandenburg*,<sup>45</sup> *Cohen*,<sup>46</sup> the Pentagon Papers case,<sup>47</sup> and the flagburning cases.<sup>48</sup> The result is something more than a collection of "greatest First Amendment hits." It is an account of the development of certain core free-speech principles: that the people are sovereign in a democracy; that wide open debate is necessary if the people are to perform their sovereign function; that government regulation of such debate should ever be distrusted. In turn, these principles provide the measure of current First Amendment problems. Thus, Lewis makes a compelling case that the greatest of all obstacles to a flourishing system of freedom of expression is governmental secrecy, especially in matters pertaining to national security (at 241-43). And indeed, this matter resonates with *Sullivan* more strongly than does the run-of-the-mine libel action.

Above all, as Lewis highlights, *Sullivan* is a statement—the Court's strongest statement—of core First Amendment values. In its substance—and also, if the two can be separated, in its rhetoric—the decision speaks of the potential of democracy, the role of free expression in realizing that potential, the corresponding threat such expression may pose to those wielding power. At the same time, the decision speaks to the widest possible audience—not to the press, as in so many of the Court's libel cases, but to the American public. It reminds us of the kind of public discourse we should aspire to, as well as of what we must tolerate to attain it. *Sullivan* goes only part of the way toward solving particular cases and problems; in the field where it has been most studiously applied, it has produced a mixed bag of consequences. But at the most general level—as a statement of enduring principle addressed to the American people—it is indeed a marvel.

40. *Abrams v. United States*, 250 U.S. 616, 624-31 (1919) (Holmes, J., dissenting); *Gitlow v. New York*, 268 U.S. 652, 672-73 (1925) (Holmes, J., dissenting); *Whitney v. California*, 274 U.S. 357, 374-77 (1927) (Brandeis, J., concurring); *United States v. Schwimmer*, 279 U.S. 644, 653-55 (1929) (Holmes, J., dissenting).

41. *Near v. Minnesota*, 283 U.S. 697 (1931).

42. *Grosjean v. American Press Co.*, 297 U.S. 233 (1936).

43. *Bridges v. California*, 314 U.S. 252 (1941).

44. *Bond v. Floyd*, 385 U.S. 116 (1966).

45. *Brandenburg v. Ohio*, 395 U.S. 444 (1968).

46. *Cohen v. California*, 403 U.S. 15 (1971).

47. *New York Times Co. v. United States*, 403 U.S. 713 (1971).

48. *Texas v. Johnson*, 491 U.S. 397 (1989); *United States v. Eichman*, 496 U.S. 310 (1990).

And here lies the ultimate value of Lewis's work as well. In *Make No Law*, Lewis sometimes fails to discriminate among different kinds of First Amendment problems, to explore as deeply as he might the range of considerations involved in particular free speech controversies. But as an expounder of broad principle, he has few, if any, peers. And perhaps it is more important that the broad audience he is addressing have a deep commitment to the principle than a subtle understanding of the ways it can be applied or a fine appreciation of its limits. Like *Sullivan* itself, Lewis's work bears more than a passing resemblance to a morality play. Which is to say that although neither tells us everything, both instruct us as to what is most important.

# American Bar Foundation

---

## BOARD OF DIRECTORS

---

*President*, Robert W. Bennett, Northwestern University School of Law • *Vice-President*, John C. Deacon, of the Arkansas Bar • *Secretary*, Robert MacCrate, of the New York Bar • *Treasurer*, John L. Carey, of the Indiana Bar

Brooksley Born, of the D.C. Bar • Kenneth J. Burns, Jr., of the Illinois Bar • Hon. Jerome Farris, United States Court of Appeals, Ninth Circuit • Daniel Gourash, of the Ohio Bar • Harry E. Groves, University of North Carolina School of Law • Martha L. Minow, Harvard University Law School • Wm. Reece Smith, Jr., of the Florida Bar • Robert A. Stein, University of Minnesota Law School • Cathy S. Wright, of the Alabama Bar

*ex officio*:

J. Michael McWilliams, President, American Bar Association • R. William Ide, III, President-Elect, American Bar Association • Philip S. Anderson, Chairman, House of Delegates, American Bar Association • John L. Carey, Treasurer, American Bar Association • Herschel H. Friday, Jr., President, American Bar Endowment • Richard L. Thies, Chair, The Fellows of the American Bar Foundation • James W. Hewitt, Vice-Chair, The Fellows of the American Bar Foundation • Marna S. Tucker, Secretary, The Fellows of the American Bar Foundation

---

## OFFICERS

---

Bryant G. Garth, *Director*  
Joanne Martin, *Assistant Director*

Anne Tatalovich, *Assistant Director*  
Hayden C. Flor, Jr., *Financial Officer*

---

## RESEARCH FELLOWS

---

Lori B. Andrews, J.D. Yale University  
Jonathan D. Casper, Ph.D. Yale University  
John L. Comaroff, Ph.D. University of London  
Stephen Daniels, Ph.D. University of Wisconsin  
Shari S. Diamond, Ph.D. Northwestern University;  
J.D. University of Chicago  
John J. Donohue, J.D. Harvard University;  
Ph.D. Yale University  
William L. F. Felstiner, J.D. Yale University  
Bryant G. Garth, J.D. Stanford University; Ph.D.  
European University Institute  
Janet A. Gilboy, Ph.D. Northwestern University  
Terence C. Halliday, Ph.D. University of Chicago  
James J. Heckman, Ph.D. Princeton  
Carol A. Heimer, Ph.D. University of Chicago  
John P. Heinz, LL.B. Yale University

Keith N. Hylton, J.D. Harvard University; Ph.D.  
Massachusetts Institute of Technology  
Wayne A. Kerstetter, J.D. University of Chicago  
Karyl A. Kinsey, Ph.D. University of Chicago  
E. Allan Lind, Ph.D. University of North  
Carolina at Chapel Hill  
Arthur F. McEvoy, J.D. Stanford University;  
Ph.D. University of California, San Diego  
Elizabeth Mertz, J.D. Northwestern University;  
Ph.D. Duke University  
Robert L. Nelson, J.D., Ph.D. Northwestern  
University  
Susan P. Shapiro, Ph.D. Yale University  
Peter Siegelman, Ph.D. Yale University  
Kent W. Smith, Ph.D. Harvard University  
Christopher L. Tomlins, Ph.D. Johns Hopkins

---

## LIAISON RESEARCH SERVICES PROGRAM

---

Joanne Martin, M.M. Northwestern University, J.D. Loyola University  
Barbara A. Curran, LL.B. University of Connecticut, LL.M. Yale University

CLINTON LIBRARY PHOTOCOPY

Reprinted from  
**Texas Law Review**

---

FOR JUSTICE MARSHALL

*Elena Kagan*

---

Copyright © 1993 Texas Law Review Association

---

MAY 1993

VOL. 71 No. 6

---

CLINTON LIBRARY PHOTOCOPY



## For Justice Marshall

Elena Kagan\*

A few days after Thurgood Marshall's death, I stood for a time at his flag-draped casket, then lying in state at the Supreme Court, and watched the people of Washington celebrate his life and mourn his passing. There would be, the next day, a memorial service for the Justice in the National Cathedral, a grand affair complete with a Bible reading by the Vice President and eulogies by the Chief Justice and other notables. That service would have its moments, but it would not honor Justice Marshall as the ordinary people of Washington did. On the day the Justice's casket lay in state, some 20,000 of them came to the Court and stood in bitter cold for upwards of an hour in a line that snaked down the Supreme Court steps, down the block, around the corner, and down the block again. The Justice's former clerks took turns standing at the casket, acting as a kind of honor guard, as these thousands of people filed by. Passing before me were people of all races, of all classes, of all ages. Many came with children and spoke, as they circled the casket, of the significance of Justice Marshall's life. Some offered tangible tributes—flowers or letters addressed to Justice Marshall or his family. One left at the side of the casket a yellowed slip opinion of *Brown v. Board of Education*.<sup>1</sup> There never before has been such an outpouring of love and respect for a Supreme Court Justice, and there never will be again. As I stood and watched, I felt (as I will always feel) proud and honored and grateful beyond all measure to have had the chance to work for this hero of American law and this extraordinary man.

I first spoke with Justice Marshall in the summer of 1986, a few months after I had applied to him for a clerkship position. (It seems odd to call him Justice Marshall in these pages. My co-clerks and I called him "Judge" or "Boss" to his face, "TM" behind his back; he called me, to my face and I imagine also behind my back, "Shorty.") He called me one day and, with little in the way of preliminaries, asked me whether I still wanted

---

\* Assistant Professor, University of Chicago Law School; law clerk to Justice Marshall, 1987 Term. A.B. 1981, Princeton University; M. Phil. 1983, Oxford University; J.D. 1986, Harvard Law School.

1. 347 U.S. 483 (1954).

a job in his chambers. I responded that I would love a job. "What's that?" he said, "you already have a job?" I tried, in every way I could, to correct his apparent misperception. I yelled, I shouted, I screamed that I did not have a job, that I wanted a job, that I would be honored to work for him. To all of which he responded: "Well, I don't know, if you already have a job . . ." Finally, he took pity on me, assured me that he had been in jest, and confirmed that I would have a job in his chambers. He asked me, as I recall, only one further question: whether I thought I would enjoy working on dissents.

So went my introduction to Justice Marshall's (sometimes wicked) sense of humor. He took constant delight in baffling and confusing his clerks, often by saying the utterly ridiculous with an air of such sobriety that he half-convincing us of his sincerity. (There was the time, for example, when he announced sadly that he would have to recuse himself from *Gwaltney of Smithfield v. Chesapeake Bay Foundation*.<sup>2</sup> When we pressed him for a reason, he hemmed and hawed for many minutes, only finally to say: "Because I l-o-o-o-o-v-e their ham." When we laughed, he assumed an attitude of great indignation and began instructing us on proper recusal policy. It was early in the Term; perhaps we may be forgiven for thinking for a moment that, after all, this was not a joke.) He had an endless supply of jokes, not all of them, I must admit, appropriate to print in the pages of a law review. And he was the greatest comic storyteller I have ever heard, or ever expect to hear. This talent, I think, may be impossible to communicate to those never exposed to it. It was a matter of timing (the drawn-out lead-up, the pregnant pause), of vocal intonations and inflections, and most of all of facial expressions (the raised brow, the sparkling eyes, the sidelong glance). Suffice it to say that at least once in the course of every meeting we had with him (and those were frequent), my co-clerks and I would find ourselves holding our sides and gasping for breath, as we struggled to regain our composure.

Thinking back, I'm not sure why we laughed so hard—or rather, I'm not sure why Justice Marshall told his stories so as to make us laugh—because most of the stories really weren't funny. To be sure, some were pure camp. (When Justice Marshall was investigating racial discrimination in the military in Korea, a soldier demanded that he provide a password; the hulking (and, of course, black) Marshall looked down at the soldier and asked, "Do you really think I'm North Korean?" And when assisting in the drafting of the Kenyan Constitution, the Justice was introduced to Prince Philip. "Do you care to hear my opinion of lawyers?" Prince Philip asked in posh British tones, mimicked to great comic effect by Justice Marshall. "Only," Justice Marshall replied—before the two discov-

---

2. 484 U.S. 49 (1987).

ered mutual ground in a taste for bourbon—"if you care to hear my opinion of princes.") But most of the stories, if told by someone else, would have expressed only sorrow and grimness. They were stories of growing up black in segregated Baltimore, subject to daily humiliation and abuse. They were stories of representing African-American defendants in criminal cases—often capital cases—in which a fair trial was not to be hoped for, let alone expected. (He knew he had an innocent client, Justice Marshall said, when the jury returned a sentence of life imprisonment, rather than execution.) They were stories of the physical danger (the lynch mobs, the bomb-throwers, the police themselves) that the Justice frequently encountered as he traversed the South battling state-imposed segregation. They were stories of prejudice, violence, hatred, fear; only as told by Justice Marshall could they ever have become stories of humor and transcending humanity.

The stories were something more than diversions (though, of course, they were that too). They were a way of showing us that, bright young legal whipper-snappers though we were, we did not know everything; indeed, we knew, when it came to matters of real importance, nothing. They were a way of showing us foreign experiences and worlds, and in doing so, of reorienting our perspectives on even what had seemed most familiar. And they served another function as well: they reminded us, as Justice Marshall thought all lawyers (and certainly all judges) should be reminded, that behind law there are stories—stories of people's lives as shaped by law, stories of people's lives as might be changed by law. Justice Marshall had little use for law as abstraction, divorced from social reality (he muttered under his breath for days about Judge Bork's remark that he wished to serve on the Court because the experience would be "an intellectual feast"); his stories kept us focused on law as a source of human well-being.

That this focus made the Justice no less a "lawyer's lawyer" should be obvious; indeed, I think, quite the opposite. I knew, of course, before I became his clerk that Justice Marshall had been the most important—and probably the greatest—lawyer of the twentieth century. I knew that he had shaped the strategy that led to *Brown v. Board of Education* and other landmark civil rights cases; that he had achieved great renown (indeed, legendary status) as a trial lawyer; that he had won twenty-nine of the thirty-two cases he argued before the Supreme Court. But in my year of clerking, I think I saw what had made him great. Even at the age of eighty, his mind was active and acute, and he was an almost instant study. Above all, though, he had the great lawyer's talent (a talent many judges do not possess) for pinpointing a case's critical fact or core issue. That trait, I think, resulted from his understanding of the pragmatic—of the way in which law worked in practice as well as on the books, of the way in

which law acted on people's lives. If a clerk wished for a year of spinning ever more refined (and ever less plausible) law-school hypotheticals, she might wish for a clerkship other than Justice Marshall's. If she thought it more important for a Justice to understand what was truly going on in a case and to respond to those realities, she belonged in Justice Marshall's chambers.

None of this meant that notions of equity governed Justice Marshall's vote in every case; indeed, he could become quite the formalist at times. During the Term I clerked, the Court heard argument in *Torres v. Oakland Scavenger Co.*<sup>3</sup> There, a number of Hispanic employees had brought suit alleging employment discrimination. The district court dismissed the suit, and the employees' lawyer filed a notice of appeal. The lawyer's secretary, however, inadvertently omitted the name of one plaintiff from the notice. The question for the Court was whether the appellate court had jurisdiction over the party whose name had been omitted; on this question rode the continued existence of the employee's discrimination claim. My co-clerks and I pleaded with Justice Marshall to vote (as Justice Brennan eventually did) that the appellate court could exercise jurisdiction. Justice Marshall refused. As always when he disagreed with us, he pointed to the framed judicial commission hanging on his office wall and asked whose name was on it. (Whenever we told Justice Marshall that he "had to" do something—join an opinion, say—the Justice would look at us coldly and announce: "There are only two things I *have to* do—stay black and die." A smarter group of clerks might have learned to avoid this unfortunate grammatical construction.) The Justice referred in our conversation to his own years of trying civil rights claims. All you could hope for, he remarked, was that a court didn't rule against you for illegitimate reasons; you couldn't hope, and you had no right to expect, that a court would bend the rules in your favor. Indeed, the Justice continued, it was the very existence of rules—along with the judiciary's felt obligation to adhere to them—that best protected unpopular parties. Contrary to some conservative critiques, Justice Marshall believed devoutly—believed in a near-mystical sense—in the rule of law. He had no trouble writing the *Torres* opinion.

Always, though, Justice Marshall believed that one kind of law—the Constitution—was special, and that the courts must interpret it in a special manner. Here, more than anywhere else, Justice Marshall allowed his personal experiences, and the knowledge of suffering and deprivation gained from those experiences, to guide him. Justice Marshall used to tell of a black railroad porter who noted that he had been in every state and every city in the country, but that he had never been anyplace where he had to

---

3. 487 U.S. 312 (1988).

put his hand in front of his face to know that he was black. Justice Marshall's deepest commitment was to ensuring that the Constitution fulfilled its promise of eradicating such entrenched inequalities—not only for African-Americans, but for all Americans alike.

The case I think Justice Marshall cared about most during the Term I clerked for him was *Kadrmas v. Dickinson Public Schools*.<sup>4</sup> The question in *Kadrmas* was whether a school district had violated the Equal Protection Clause by imposing a fee for school bus service and then refusing to waive the fee for an indigent child who lived sixteen miles from the nearest school. I remember, in our initial discussion of the case, opining to Justice Marshall that it would be difficult to find in favor of the child, Sarita Kadrmas, under equal protection law. After all, I said, indigency was not a suspect class; education was not a fundamental right; thus, a rational basis test should apply, and the school district had a rational basis for the contested action. Justice Marshall (I must digress here) didn't always call me "Shorty"; when I said or did something particularly foolish, he called me (as, I hasten to add, he called all his clerks in such situations) "Knucklehead." The day I first spoke to him about *Kadrmas* was definitely a "Knucklehead" day. (As I recall, my handling of *Kadrmas* earned me that appellation several more times, as Justice Marshall returned to me successive drafts of the dissenting opinion for failing to express—or for failing to express in a properly pungent tone—his understanding of the case.) To Justice Marshall, the notion that government would act so as to deprive poor children of an education—of "an opportunity to improve their status and better their lives"<sup>5</sup>—was anathema. And the notion that the Court would allow such action was even more so; to do this would be to abdicate the judiciary's most important responsibility and its most precious function.

For in Justice Marshall's view, constitutional interpretation demanded, above all else, one thing from the courts: it demanded that the courts show a special solicitude for the despised and disadvantaged. It was the role of the courts, in interpreting the Constitution, to protect the people who went unprotected by every other organ of government—to safeguard the interests of people who had no other champion. The Court existed primarily to fulfill this mission. (Indeed, I think if Justice Marshall had had his way, cases like *Kadrmas* would have been the only cases the Supreme Court heard. He once came back from conference and told us sadly that the other Justices had rejected his proposal for a new Supreme Court rule. "What was the rule, Judge?" we asked. "When one corporate fat cat sues another corporate fat cat," he replied, "this Court shall have no juris-

---

4. 487 U.S. 450 (1988).

5. *Id.* at 468-69 (Marshall, J., dissenting).

diction.") The nine Justices sat, to put the matter baldly, to ensure that Sarita Kadrmas could go to school each morning. At any rate, this was why they sat in Justice Marshall's vision of the Court and Constitution. And however much some recent Justices have sniped at that vision, it remains a thing of glory.

During the year that marked the bicentennial of the Constitution, Justice Marshall gave a characteristically candid speech. He declared that the Constitution, as originally drafted and conceived, was "defective"; only over the course of 200 years had the nation "attain[ed] the system of constitutional government, and its respect for . . . individual freedoms and human rights, we hold as fundamental today."<sup>6</sup> The Constitution today, the Justice continued, contains a great deal to be proud of. "[B]ut the credit does not belong to the Framers. It belongs to those who refused to acquiesce in outdated notions of 'liberty,' 'justice,' and 'equality,' and who strived to better them."<sup>7</sup> The credit, in other words, belongs to people like Justice Marshall. As the many thousands who waited on the Supreme Court steps well knew, our modern Constitution is his.

---

6. Thurgood Marshall, *The Constitution's Bicentennial: Commemorating the Wrong Document?*, 40 VAND. L. REV. 1337, 1338 (1987).

7. *Id.* at 1341.