Florida Law Review

Volume 57 | Issue 4 Article 2

September 2005

Containing Unprotected Speech

Heidi Kitrosser

CONTAINING UNPROTECTED SPEECH

Heidi Kitrosser[*]

---

* Assistant Professor, Brooklyn Law School; visiting Associate Professor, University of Minnesota Law School. This Article benefited from very helpful discussions with Rick Brooks, Andy Koppelman, and Marty Redish during the Article's formulative, predrafting period. This Article also benefited from very helpful comments on early drafts by Larry Alexander, Jenny Diamond Cheng, Rick Garnett, Andy Koppelman, Marty Redish, Dana Brakman Reiser, Larry Solan, Jim Speta, Nelson Tebbe, Kim Yuracko, and participants in workshops at the 2004 Law & Society conference, Brooklyn Law School, the New York Junior Faculty Colloquium at Fordham Law School, Rutgers Law School—Newark, the University of San Diego and the University of Minnesota. Kim Brown and Jason Litwack, both of the Brooklyn Law School class of '06, provided excellent research assistance. This Article was supported by a summer research grant from Brooklyn Law School.

I. INTRODUCTION

The Supreme Court long has deemed a few categories of speech so harmful and so lacking in value[1] as to be unworthy of First

1. Because neither history nor text offer much insight regarding the meaning and scope of the First Amendment's free speech protections, scholars and courts tend to rely on theories of free speech value. That is, they rely on varying theories about the value that inheres in free speech to deduce the scope of free speech protections. *See generally* ERWIN CHEMERINSKY, CONSTITUTIONAL LAW: PRINCIPALS AND POLICIES § 11.1.1 (2d ed. 2002) (discussing the difficulty in ascertaining the framers' intent regarding the meaning of the First Amendment); Martin H. Redish, *The Value of Free Speech*, 130 U. PA. L. REV. 591, 591, 595 n.27 (1982) [hereinafter Redish, *Value of Free Speech*] (discussing the lack of a unified theory of free speech value). Theories of free speech value

Amendment protection.[2] Under this approach, which this Article calls categorization doctrine, legislatures may regulate—even ban—unprotected speech categories in their entirety.[3] Specifically, legislatures may pass laws punishing all speech within the following categories: threats,[4] fighting words,[5] obscenity,[6] child pornography,[7] and speech that imminently incites illegal activity.[8]

Because of categorization doctrine, one might reasonably assume that legislatures can choose either to regulate all speech within an unprotected speech category, e.g., by imposing a ban on all threats, or instead to regulate subcategories of such speech, e.g., by banning only race-based threats or threats that relate to terrorism. This assumption could stem from a simple "greater includes lesser" rationale to the effect that the power to regulate all threats must include the power to regulate some threats. Or it could stem from the notion that the rationale that supports making a category of speech "unprotected" supports making any subcategory of that larger category "unprotected." Hence, if all threats are so lacking in value or so harmful as to justify the withholding of First Amendment protection, surely the same can be said of race-based threats or threats that relate to terrorism.

These assumptions are as wrong descriptively as they seem reasonable normatively. They are wrong descriptively because the Supreme Court has made clear for over a decade that while legislatures constitutionally may regulate unprotected speech categories in their entirety, it raises serious constitutional concerns when legislatures single out only "content-based"

---

abound and can be divided roughly into "intrinsic" and "instrumental" theories. *See, e.g.*, ZECHARIAH CHAFEE JR., FREE SPEECH IN THE UNITED STATES 33 (1954); Thomas I. Emerson, *Toward a General Theory of the First Amendment*, 72 YALE L.J. 877, 878-79 (1963); Heidi Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows: Communicative Manner and the First Amendment*, 96 NW. U. L. REV. 1339, 1373-74 (2002) [hereinafter Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows*]. Intrinsic theories emphasize the intrinsic good of free speech to the individual speaker, generally for reasons relating to dignity, self-fulfillment, or self-realization. Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows*, *supra*, at 1374. Instrumental theories emphasize the instrumental role that free speech can play in improving society, either as a general matter or in terms of its democratic governance processes. *Id.*

2. *See, e.g.*, Virginia v. Black, 538 U.S. 343, 358-59 (2003) (describing categories of unprotected speech).

3. *Id.*

4. *See, e.g.*, Watts v. United States, 394 U.S. 705, 707 (1969); Planned Parenthood v. Am. Coalition of Life Activists, 290 F.3d 1058, 1062-63 (9th Cir. 2002) (en banc).

5. *See, e.g.*, Texas v. Johnson, 491 U.S. 397, 409 (1989); Chaplinsky v. New Hampshire, 315 U.S. 568, 571-72 (1942).

6. *See, e.g.*, Miller v. California, 413 U.S. 15, 23-24 (1973).

7. *See, e.g.*, New York v. Ferber, 458 U.S. 747, 756-63 (1982).

8. *See, e.g.*, Brandenburg v. Ohio, 395 U.S. 444, 447 (1969).

subcategories of unprotected speech.[9] In other words, serious constitutional concerns arise when legislatures single out subcategories of unprotected speech defined by viewpoint, subject matter, or communicative impact.[10] Thus, while it is constitutionally unproblematic for legislatures to ban all threats, serious constitutional problems arise when legislatures ban only those threats that are based on race or that relate to terrorism. While such restrictions are not per se unconstitutional, they must pass challenging doctrinal tests to be upheld.[11]

Despite the momentousness of the Supreme Court's conclusions that unprotected speech is not fully unprotected and that content-based regulations of such speech must pass tough doctrinal tests, the Court offers only the barest of rationales for these conclusions. Specifically, the Court suggests that content-based regulations of unprotected speech are intrinsically troubling and that they cause instrumental problems by skewing debate in the marketplace of ideas.[12] Put differently, the Court applies the rationale for the content-distinction rule, whereby content-based regulations of *protected* speech are deemed presumptively unconstitutional for intrinsic and instrumental reasons, to the realm of *unprotected* speech.[13] The problem with such application is that the content-distinction rule is based largely on the premise that protected speech is of significant intrinsic and instrumental value.[14] The rule does not translate to the realm of unprotected speech if a major premise of categorization doctrine is taken seriously. This premise is that unprotected

---

9. *See, e.g.*, R.A.V. v. City of St. Paul, 505 U.S. 377, 383-84 (1992).

10. *Compare id.* at 391 (deeming the St. Paul ordinance problematic because it regulates based on subject matter and possibly viewpoint), *with* Virginia v. Black, 538 U.S. 343, 360-63 (2003) (subjecting a cross-burning statute to *R.A.V.*'s standards because it regulates based on the use of a particular communicative symbol). Of course, the definition of "content" is far from clear, although it typically is deemed to include at least viewpoint and subject matter, and sometimes also to include any elements—such as the use of communicative symbols—relating to speech's "communicative impact." *See* Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows*, *supra* note 1, at 1340.

11. *R.A.V.*, 505 U.S. at 387-89, 395-96 (explaining standards applicable to content-based regulations of unprotected speech).

12. *Id.* at 384-90.

13. *Id.* at 387 (referring to a major rationale for the general content-distinction rule); *see also* Alan E. Brownstein, *Rules of Engagement for Cultural Wars: Regulating Conduct, Unprotected Speech, and Protected Expression in Anti-Abortion Protests*, 29 U.C. DAVIS L. REV. 553, 589-611 (1996) (offering justifications for the general content-distinction rule); Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows, supra* note 1, at 1392-96 (same); Geoffrey R. Stone, *Content Regulation and the First Amendment*, 25 WM. & MARY L. REV. 189, 200-33 (1983) (same).

14. In other words, because protected speech has intrinsic value, there is something intrinsically troubling about singling it out for regulation based on its content. And because protected speech is instrumentally beneficial to the marketplace of ideas, it makes sense to worry that content-based regulations will skew debate in the marketplace. *See, e.g.*, Brownstein, *supra* note 13, at 589-624.

speech categories are so intrinsically valueless and so instrumentally harmful that their abolition is consistent with, possibly even dictated by, free speech values. From this premise, it simply does not follow that the uneven regulation of unprotected speech is intrinsically or instrumentally troubling.[15]

The doctrinal standards that the Court applies to content-based regulations of unprotected speech suffer badly for their lack of a coherent rationale. The Court has crafted a grab bag of standards that range from the superfluous to the affirmatively damaging. For example, the Court purports to apply strict scrutiny to content-based regulations of unprotected speech when in fact it applies a heightened, virtually insurmountable level of scrutiny.[16] The Court also crafts exceptions to its presumption against such regulations, but its most promising exception morphs in character from case to case, making it subject to manipulation depending on the Court's gut feeling about a given regulation.[17] Finally, the Court hints that it holds viewpoint-based regulations in particular disdain and that it is less concerned about regulations based on the use of particular symbols or expressive modes.[18] The doctrinal implications of this hint are not entirely clear, which creates another avenue for *ad hocery* based on the Court's hunches about particular regulations. To the extent that this hint does manifest itself in doctrine, this is bad news for free speech, as it enables content-based regulations to pass muster so long as they can be described, however questionably, as based on communicative impact rather than viewpoint or subject matter.[19]

This Article seeks to fill the gap left by existing doctrine and scholarship by establishing a sound theoretical footing for the position that content-based regulations of unprotected speech raise important First Amendment concerns. Once such footing is established, coherent doctrinal standards can be created. This Article thus establishes a solid theoretical basis for the position that content-based regulations of unprotected speech raise First Amendment concerns. The Article then crafts new doctrinal standards for such regulations and explains the practical implications of these standards for legislation.

As a matter of theoretical grounding, this Article argues that the missing analytical link in doctrine and scholarship can be found in the close relationship between protected speech and unprotected speech. From this close relationship stems a phenomenon that this Article calls the

15. *See id.* at 611–24.
16. *See infra* Part V.B.1.
17. *See infra* Part V.B.1-2.
18. *See infra* Part V.B.1-2.
19. See *infra* notes 283-85 and accompanying text (describing how this risk came partly to fruition in the Supreme Court's reasoning and the argument of the state in *Virginia v. Black*).

categorization paradox: the notion that categorization is both necessary to free speech doctrine and necessarily problematic. Categorization is necessary because it is a relatively speech-protective means of accounting, doctrinally, for the fact that speech sometimes must be regulated for its content. At the same time, categorization is necessarily problematic because the same qualities that might make speech unprotected can bring the same speech very close on the doctrinal spectrum to protected, even deeply valued, speech. Given this inescapably close relationship between protected and unprotected speech, the reach and use of unprotected speech categories must be carefully contained. The categorization paradox gives rise, in short, to what this Article calls the containment principle.

The containment principle embodies the notion that content-based regulations of unprotected speech must be contained to defend against two risks. First, such regulations operate within a realm of at least some free speech value and thus raise free speech value-based concerns similar to those that underscore the content-distinction rule. In particular, such regulations make it easier for legislatures to regulate speech, enabling them to hone in solely on the most unpopular forms of unprotected speech. By so acting, legislatures may skew public discourse, cause intrinsic harms to speakers, and cause intrinsic social harms borne of legislative bad faith. Second, such regulations threaten to serve as vehicles to infringe upon unprotected speech's close cousin: protected speech.

Identification of the second risk—that content-based regulations of unprotected speech will infringe on protected speech—is of particular importance, both because of its novelty and because of the strong light it sheds on the close relationship between protected speech and unprotected speech. The risk takes three forms. First, such regulations can cause any "chilling effect"[20] on protected speech to be distributed in an uneven, content-based manner. Second, such regulations threaten to punish persons partly for using protected speech and partly for using unprotected speech. For example, a regulation punishing race-based threats may be drawn or applied in such a way as to punish protected, racist thoughts or speech that accompany, but that do not form part of, the relevant threats.[21] Third, such regulations threaten to steer speakers and factfinders toward the conclusion that speech that might otherwise be deemed merely close to the "unprotected line" in fact crosses the line where a targeted content element

20. "Chilling effect," a term frequently invoked in free speech doctrine and scholarship, refers to the risk that speech restrictions will inhibit more speech than they technically prohibit because speakers fearful of punishment will err on the side of caution. *See, e.g.*, CHEMERINSKY, *supra* note 1, § 11.3.5.2 ("[E]xpression . . . can be chilled and limited by tort liability.").

21. The risk exists, in short, when a regulation insufficiently connects the targeted content element to the unprotected speech category, thus defining punishable expression partly by an unprotected element and partly by a protected speech element.

is present.[22] For example, a regulation on threats conveyed through cross-burning may serve as a cue to factfinders that a cross-burning is inherently likely to constitute a threat, given the relative salience of the cross-burning factor and the relative ambiguity as to what constitutes a punishable threat.[23]

After establishing these theoretical foundations, this Article explains that the grab bag of doctrinal standards crafted by the Court does not mesh well with these foundations. This Article crafts new doctrinal standards responsive to this Article's theoretical foundations and explains that these standards offer more satisfying and more predictable approaches to statutes such as those cited above.[24] Specifically, this Article explains that only two doctrinal standards are called for in assessing content-based regulations of unprotected speech. First, courts should ask whether a content-based regulation of unprotected speech relates substantially to a harm of the same nature, but more compelling, than that against which the larger unprotected speech category is directed. This question—which resembles, but is less malleable than, one of the exceptions crafted by the Supreme Court[25]—is designed to ensure that a regulation is important enough to outweigh any containment-based risks, and that the regulation is sufficiently related to the larger unprotected speech category to affect only unprotected speech. Second, courts also should ask if a regulation threatens protected speech in a manner not likely to be identified through a general standard.[26] Finally, this Article makes clear that a hierarchy of "better" and "worse" content-based regulations—specifically, a hierarchy of viewpoint-based, subject-matter-based, and communicative impact-based regulations—is not appropriate. While the hierarchy has intuitive

---

22. In other words, a targeted content element may serve as a heuristic device, suggesting to speakers and factfinders that the thin line between protected speech and unprotected speech is bridged where the targeted content element is present.

23. Similarly, a regulation on threats conveyed through cross-burning may create a public perception that cross-burning is prohibited per se. *See infra* notes 155-60 and accompanying text (explaining that this risk manifests itself in the facts underlying *Black* and that Justice Souter's concurrence makes reference to this risk).

24. *See supra* notes 21, 23 and accompanying text.

25. *See infra* notes 219-27 and accompanying text.

26. The need for this standard is exemplified by the facts of *Virginia v. Black*, in which a statutory provision making cross-burning prima facie evidence that a threat had been made suggested a legislative desire to skew factfinder decisions toward the punishment of even nonthreatening cross-burnings. *See infra* note 286 and accompanying text. Additionally, while this standard necessarily suffers from some malleability, it is preferable to err on the side of speech-protective malleability than to err on the side of speech-restrictive malleability in using a catch-all test. The Supreme Court does the latter in asking, as one of its grab bag of standards, whether something about a regulation under review suggests that the regulation poses no threat that the "official suppression of ideas" is afoot. *See infra* notes 211-13 and accompanying text.

appeal, it is not responsive to the concerns posed by content-based regulations of unprotected speech.

The present time is a uniquely fitting one in which to reconsider this longstanding conundrum of free speech law. First, neither the Court nor commentators have considered the implications of the protected speech/unprotected speech relationship for content-based regulations of unprotected speech. In particular, neither the Court nor commentators have considered the threat that content-based regulations of unprotected speech pose to protected speech. Second, a revisiting of the Court's doctrine is particularly appropriate now. While the Court first formulated its doctrine in the 1992 case of *R.A.V. v. City of St. Paul*,[27] the Court recently revisited the doctrine in the 2003 case of *Virginia v. Black*.[28] While the *Black* Court purported to embrace the doctrine and theoretical underpinnings of *R.A.V.*, it applied *R.A.V.*'s doctrinal standards in a way that some deem a sharp break from *R.A.V.*[29] *Black* thus sheds important new light on the doctrinal standards crafted in *R.A.V.* Among other things, it suggests that the standards are troublingly malleable.[30] It also bolsters the view that the Court places substantial weight on its characterization of a regulation as viewpoint-based or communicative impact-based.[31] These points cast light, in turn, on the poverty of *R.A.V.*'s theoretical underpinnings and their consequent inability to generate useful doctrinal standards.

In short, the time is ripe to revisit the doctrine and theory of *R.A.V.* and its progeny. This Article begins this task in Parts II and III by describing and critiquing existing case law and scholarship. Part II describes and critiques the theoretical foundations of *R.A.V.* and *Black*. Part III evaluates and critiques existing scholarship on the topic, most of which was written in the wake of *R.A.V.* Part III explains that existing scholarship neither reassesses the role and scope of categorization doctrine nor looks beyond the assumption that the regulations at issue impact only unprotected speech.

Parts IV and V develop this Article's alternative theoretical and doctrinal approaches. Part IV develops the former, explaining the categorization paradox and the containment principle as new means to explain why content-based regulations of unprotected speech raise constitutional concerns. Part V develops the latter, explaining the doctrinal

---

27. 505 U.S. 377 (1992).

28. 538 U.S. 343 (2003).

29. *See* Guy-Uriel E. Charles, *Colored Speech: Cross Burnings, Epistemics, and the Triumph of the Crits?*, 93 GEO. L.J. 575, 575-78, 603-07 (2005); Frederick Schauer, *Intentions, Conventions, and the First Amendment: The Case of Cross-Burning*, 2003 SUP. CT. REV. 197, 204-09.

30. *See infra* text accompanying notes 258-59.

31. *See infra* note 246 and text accompanying notes 260-61.

standards that should apply to content-based regulations of unprotected speech in light of the concerns that those regulations raise.

Finally, Part VI applies this Article's approach to the regulations at issue in *R.A.V.* and *Black* and to two hypothetical statutes. The hypothetical statutes include a statute punishing race-based threats and a statute punishing fighting words that involve profane language. As Part VI demonstrates, judicial review of these hypothetical statutes would likely lead to different outcomes depending on whether or not this Article's approach were followed. Furthermore, even where practical outcomes would not differ, judicial reasoning would differ substantially with significant import for future cases and for free speech doctrine generally.

II. CATEGORIZATION AND THE THEORETICAL FOUNDATIONS OF *R.A.V.*

Subpart A describes existing categorization doctrine. Subpart B critiques the theoretical foundations of the doctrine crafted by the *R.A.V.* Court. Subpart B explains that the *R.A.V.* Court missed an important opportunity to recharacterize the nature and scope of categorization doctrine in a meaningful way. Because the Court missed this important analytical step, its doctrine ultimately rests on an illogical foundation.

A. *Descriptive Analysis of Categorization Doctrine*

Categorization doctrine encompasses more than categories of "fully protected" and "fully unprotected" speech. Categories of speech deemed fully unprotected include speech that intentionally incites imminent illegal action, fighting words, threats, obscenity, and child pornography.[32] While the basic meaning, if not the precise definition, of most of these categories is likely intuitively clear, the basic meaning of "fighting words" might be less obvious. It thus bears brief mention that the constitutional definition of fighting words, while increasingly controversial, refers at the very least to those words "reasonably considered direct personal insults likely to invite physical confrontation."[33] In addition to these fully unprotected speech categories, categorization doctrine also affords certain speech categories a lesser degree of protection than fully protected speech. For

32. *See supra* notes 4-8 and accompanying text. Defamation, too, generally is deemed unprotected. *See, e.g.*, GEOFFREY R. STONE ET AL., THE FIRST AMENDMENT 132-53 (1999). Defamation doctrine is relatively complicated, however, as it encompasses varying, context-specific tests designed to be particularly speech-protective in certain situations, for example, where speech concerning public figures is at issue. *See, e.g.*, *id.* Defamation doctrine is discussed *infra* at notes 173-78 and accompanying text.

33. Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows*, *supra* note 1, at 1347 n.36; *see also* Stephen W. Gard, *Fighting Words as Free Speech*, 58 WASH. U. L.Q. 531, 534 (1980).

example, commercial speech long has been deemed protected but to a somewhat lesser degree than other forms of speech.[34] And the Supreme Court often refers to speech that is indecent, but not obscene, as falling outside of the First Amendment's "core," although the doctrinal implications of this observation are far from clear.[35] For purposes of this Article, the only element of categorization doctrine considered, either in its own right or in terms of its interaction with content-distinction principles, is the doctrine's identification of certain speech categories as fully unprotected. The propriety of "less protected" speech categories and the intersection of such categories with content-distinction principles raise distinct issues that extend beyond this Article's scope.

A review of the Supreme Court's major unprotected speech cases suggests that the Court's justifications for the existence of unprotected speech can be grouped into two major rationales.[36] First, in at least one case (that of obscenity), the Court places very little emphasis on the purported harms of the speech, focusing instead on the purported valuelessness of the speech and vaguely alluding to expressive harms such as "patent[] offensive[ness]."[37] Second, and more typically, the Court focuses primarily on harm, concluding that a speech category's harms far outweigh any expressive value that might exist.[38] To varying degrees, this justification factors into the Court's analysis regarding all unprotected speech categories aside from obscenity. The Court's justifications for child pornography, for example, provide an interesting contrast to its justifications for obscenity. Whereas in the latter category the Court says

---

34. *See, e.g.*, Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n, 447 U.S. 557, 561-63 (1980).

35. *See, e.g.*, Kitrosser, *From Marshall McLuhan to Anthropomorphic* Cows, *supra* note 1, at 1346 & nn.26–27 (noting that the Supreme Court has applied strict scrutiny to restrictions upon both "core" political speech and speech which "'lie[s] at the periphery of First Amendment concern'") (quoting FCC v. Pacifica Found., 438 U.S. 726, 743 (1978)).

36. The justification for the existence of unprotected speech categories frequently is linked to dicta from the seminal fighting words case of *Chaplinsky v. New Hampshire*, 315 U.S. 568 (1942). Specifically, *Chaplinsky* often is cited for the proposition that

> [t]here are certain well-defined and narrowly limited classes of speech, the prevention and punishment of which has never been thought to raise any Constitutional problem. . . . It has been well observed that such utterances are no essential part of any exposition of ideas, and 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.

Id. at 571-72; *see also* Virginia v. Black, 538 U.S. 343, 358 (2003) (quoting *Chaplinsky* dicta); R.A.V. v. City of St. Paul, 505 U.S. 377, 382-83 (1992) (same).

37. *See* Miller v. California, 413 U.S. 15, 23–30 (1973).

38. New York v. Ferber, 458 U.S. 747, 763-64 (1982).

very little about harm but deems obscenity lacking in "serious literary, artistic, political, or scientific value,"[39] in the former category the Court acknowledges that child pornography may on rare occasions have value, but it concludes that the harms to its juvenile subjects far outweigh any value.[40] An emphasis on overwhelming harms also is evinced in the Court's opinions regarding speech likely to incite imminent illegal activity, fighting words, and threats. In these cases, the Court indicates that speech that imminently threatens to spark a fight, that otherwise threatens to incite violence, or that achieves its effect by creating fear of physical violence bears harms sufficient to justify its suppression.[41]

The bulk of categorization cases thus suggest that unprotected speech regulation simply "raise[s no] Constitutional problem" because any speech value is de minimis and is outweighed by overwhelming harms.[42] Categorization cases also suggest that harms caused by unprotected speech not only outweigh any de minimis value but also operate in a manner antithetical to free speech value. Modern incitement doctrine, for example, seems grounded not only in the notion that speech inciting imminent lawless action causes grave harms, but also in the notion that such speech circumvents the conditions assumed in classical instrumental theories of free speech value by achieving its effect before other speech can compete against it in the marketplace of ideas.[43] And the recent, controversial split by an en banc panel of the Ninth Circuit over the constitutional definition of "true threats" was motivated partly by competing conclusions as to

---

39. *Miller*, 413 U.S. at 24.

40. *See Ferber*, 458 U.S. at 761-65; *see also* Osborne v. Ohio, 495 U.S. 103, 108 (1990) (distinguishing a statute prohibiting the private possession of child pornography from a statute prohibiting the private possession of obscenity on the basis that the state interests underlying the former "far exceed" those underlying the latter).

41. *See, e.g.*, Brandenburg v. Ohio, 395 U.S. 444, 447 (1969); Watts v. United States, 394 U.S. 705, 707 (1969); *Chaplinsky*, 315 U.S. at 571-73; *cf.* Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting) (stating that suppression on the basis that speech poses a clear and present danger is justified only when speech "so imminently threaten[s] immediate interference with the lawful and pressing purposes of the law that an immediate check is required to save the country").

42. *See infra* note 36 (quoting *Chaplinsky*, 315 U.S. at 571-72).

43. Indeed, the relatively rigid and speech-protective requirements announced in *Brandenburg v. Ohio*, 395 U.S. at 447, particularly the imminence requirement, *id.*, seem responsive to earlier dissents by Justice Holmes to the effect that speech should be left to compete in the marketplace of ideas unless speech "so imminently threaten[s] immediate interference with the lawful and pressing purposes of the law that an immediate check is required to save the country." *Abrams*, 250 U.S. at 630; *see also* Gitlow v. New York, 268 U.S. 652, 673 (1925) (Holmes, J., dissenting) ("If in the long run the beliefs expressed in proletarian dictatorship are destined to be accepted by the dominant forces of the community, the only meaning of free speech is that they should be given their chance and have their way."); FREDERICK SCHAUER, FREE SPEECH: A PHILOSOPHICAL ENQUIRY 30 (1982).

when threats circumvent the preconditions of intrinsic and instrumental free speech theories by undermining threat recipients' autonomy and ability to respond to the threats.[44] Finally, as will be discussed in Part IV, it bears noting that there is some recognition in categorization case law—particularly in the area of defamation—of the fine line separating protected and unprotected speech.[45] With a notable exception,[46] however, this recognition manifests itself only in the view that unprotected speech categories must be carefully defined. Any cautionary notes in the case law thus are reconciled with the absolutist language of most categorization cases through the implicit assumption that unprotected speech categories, *once defined*, are safe regulatory tools. It thus is assumed implicitly that unprotected speech categories, once selected and defined, can be regulated without risk to free speech value in general or to protected speech in particular.

B. R.A.V. *and Categorization Doctrine*

Section 1 begins with a brief description of *R.A.V.*'s facts and the theoretical foundations outlined in *R.A.V.* for the proposition that content-based regulations of unprotected speech are constitutionally problematic. Section 2 critiques *R.A.V.*'s theoretical foundations, emphasizing their inconsistency with categorization doctrine and the Court's failure to meaningfully reassess categorization doctrine to explain its departure from the same.

1. Descriptive Analysis of *R.A.V.*'s Background and Theoretical Foundations

The events giving rise to *R.A.V. v. City of St. Paul* began when several teenagers, including R.A.V.,[47] burned a cross inside of the fenced yard of an African-American family during predawn hours.[48] The cross-burning could have been prosecuted under a number of state statutes, including statutes punishing terroristic threats and statutes punishing property

---

44. Planned Parenthood v. Am. Coalition of Life Activists, 290 F.3d 1058, 1085-86 (9th Cir. 2002); *id.* at 1088-89 (Reinhardt, J., dissenting); *id.* at 1098-1100 (Kozinski, J., dissenting); *id.* 1101-07 (Berzon, J., dissenting).

45. *See infra* Part IV.D.

46. For discussion of the "notable exception," see *infra* notes 179-86 and accompanying text (citing Bantam Books v. Sullivan, 372 U.S. 58 (1963)).

47. R.A.V. was a juvenile at the time that he was charged with the crime. R.A.V. v. City of St. Paul, 505 U.S. 377, 380 (1992).

48. *Id.* at 379. These facts are recounted in greater detail elsewhere. *See, e.g.*, Charles R. Lawrence, III, *Crossburning and the Sound of Silence: Antisubordination Theory and the First Amendment*, 37 VILL. L. REV. 787, 787-88 (1992).

damage.[49] Instead, R.A.V. was prosecuted under a state statute prohibiting racially-motivated assaults and a local ordinance.[50] Although R.A.V. did not challenge the state assault statute under which he was prosecuted,[51] he did challenge the constitutionality of the local ordinance.[52]

The local ordinance under which R.A.V. was prosecuted provided as follows:

> Whoever places 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 commits disorderly conduct and shall be guilty of a misdemeanor.[53]

R.A.V. challenged the ordinance as overbroad, arguing that it reached well beyond the realm of unprotected speech categories—particularly fighting words—to punish protected speech.[54] The Minnesota Supreme Court disagreed, construing the ordinance to reach only fighting words[55] and deemed the ordinance constitutional.[56] The Supreme Court granted certiorari.[57]

While all nine members of the Supreme Court agreed that the local ordinance was unconstitutional, the Court divided sharply over the basis for this conclusion. Justices White, Blackmun, O'Connor, and Stevens would have voided the ordinance "by holding . . . that [it] is fatally overbroad because it criminalizes not only unprotected expression, but also expression protected by the First Amendment."[58] A majority of the Court, comprised of Justices Scalia, Rehnquist, Kennedy, Souter, and Thomas ("the *R.A.V.* Court"), accepted the Minnesota Supreme Court's construction of the ordinance as reaching only unprotected fighting words,[59] but it deemed the ordinance impermissibly content-based because

---

49. *R.A.V.*, 505 U.S. at 380 n.1.
50. *Id.* at 379-80, 380 & n.2.
51. *Id.* at 380 n.2.
52. *Id.* at 380.
53. *Id.* at 380 (quoting ST. PAUL, MINN. LEGIS. CODE § 292.02 (1990)).
54. *Id.* at 380; *see also In re* Welfare of R.A.V., 464 N.W.2d 507, 508-09 (Minn. 1991).
55. *R.A.V.*, 505 U.S. at 380-81; *Welfare of R.A.V.*, 464 N.W.2d at 510.
56. *R.A.V.*, 505 U.S. at 380-81; *Welfare of R.A.V.*, 464 N.W.2d at 510-11.
57. 501 U.S. 1204 (1991).
58. *R.A.V.*, 505 U.S. at 397 (White, J., concurring in the judgment).
59. *R.A.V.*, 505 U.S. at 381, 391. It is somewhat unclear whether the Court actually accepted the Minnesota Supreme Court's construction of the statute or whether the Court made this assumption arguendo. *See id.*

it reached only those fighting words used to "express views on disfavored subjects."[60]

The *R.A.V.* Court explained that the Court's many references over the years to unprotected speech categories falling entirely outside of "'the area of constitutionally protected speech'" are not literally true, and that content-distinction principles are applicable even to unprotected speech.[61] The *R.A.V.* Court further explained that while unprotected speech categories always can be regulated in their entirety, subcategories of such categories generally may not be singled out and regulated on content-specific bases.[62] Thus, while states and localities may regulate all fighting words without constitutional difficulty, states and localities generally may not regulate only those fighting words based on race, gender, or other specified topics.[63]

It was incumbent upon the *R.A.V.* Court to explain why content-distinction principles are relevant to unprotected speech. The Court's analysis is far from clear on this point, but it suggests two major bases for the Court's invoking content-distinction principles: instrumental concerns and intrinsic concerns. As for the former, the Court states that while unprotected speech categories such as fighting words constitute "'no *essential* part of any exposition of ideas,'"[64] the categories can be "quite expressive indeed."[65] The Court also states that a state or locality may not "license one side of a debate to fight freestyle [by allowing them to use fighting words or other unprotected speech forms], while requiring the other to follow Marquis of Queensberry rules [by prohibiting them from using fighting words or other unprotected speech forms]."[66] Taken together, these points suggest a concern by the *R.A.V.* Court that content-based regulations of unprotected speech will skew debate in the marketplace of ideas by permitting only some speakers to use unprotected speech forms. As for intrinsic concerns, the Court refers repeatedly to the general risks of content-based decisionmaking, such as the risk that the "official suppression of ideas is afoot,"[67] or the risk that regulation is based on "hostility—or favoritism—towards the underlying message expressed."[68] Such statements suggest a belief that government regulation

60. *Id.* at 391.

61. *Id.* at 383-84 (quoting Roth v. United States, 354 U.S. 476, 483 (1957)).

62. *Id.* at 383-84.

63. *Id.* at 386.

64. *Id.* at 385 (quoting Chaplinsky v. New Hampshire, 315 U.S. 568, 572 (1942) (emphasis added by *R.A.V.* Court)).

65. *Id.*

66. *Id.* at 392.

67. *Id.* at 390.

68. *Id.* at 386; *see also id.* at 388 (referring to the "danger of idea or viewpoint discrimination").

based on favoritism toward the ideas expressed by regulated behavior is intrinsically problematic.

Part V of this Article describes and critiques the doctrinal standards drawn by the *R.A.V.* Court. It suffices for now to note that the Court drew relatively challenging standards in response to the intrinsic and instrumental concerns that it outlined.[69] While the standards are not irretrievably far from the mark when viewed from a containment-based approach, they are confusing and imprecise as means of responding to the close relationship between protected and unprotected speech and the risk that the regulations at issue will infringe on protected speech. Such incoherence and imprecision are not surprising because the *R.A.V.* Court's cited concerns, which constitute its theoretical foundations, do not meaningfully address the categorization paradox and consequent containment principles. The next Section critiques *R.A.V.*'s theoretical foundations.

2. Critique of *R.A.V.*'s Theoretical Foundations

The *R.A.V.* Court's invocation of intrinsic and instrumental concerns makes sense only if one understands that free speech value is not absent from the realm of unprotected speech and that unprotected speech categories threaten protected speech. This Section explains that while the *R.A.V.* Court's analysis approaches the first observation, the analysis ultimately falters in its approach and barely addresses the second observation. As a result, the Court fails to resolve a basic inconsistency with categorization doctrine.

The *R.A.V.* Court approaches the insight of the categorization paradox in its observation that unprotected speech categories can be "quite expressive indeed"[70] and that their regulation thus might skew the marketplace of ideas. The problem with this explanation, however, is that it does not go nearly far enough. It long has been clear that unprotected speech can be regulated precisely *because* of its expressive content.[71] Indeed, the concept of unprotected speech raises so many difficulties precisely because the status of most unprotected speech categories stems from harms linked to expression. Incitement doctrine, for example, took decades to develop and remains controversial precisely because speech that intentionally incites imminent illegal activity may involve impassioned calls to action grounded in social or political views.[72] Where

69. *See infra* Part V.B.1.

70. *R.A.V.*, 505 U.S. at 384-85 (citing Chaplinsky v. New Hampshire, 315 U.S. 568, 572 (1942)).

71. *See infra* Part II.A.

72. For a brief discussion of incitement doctrine's long and tortured development, see, e.g.,

such expression is so closely "brigaded with action"[73] that it intentionally and imminently incites illegal activity, it is said to cross a fine line separating punishable triggers to action from protected expression that advocates illegal activity more abstractly.[74] While speech on the former side of the line may well generate harms so overwhelming and unstoppable as to justify punishment, it would be folly to pretend that such speech does not create its harms through expressive effect, just as speech on the latter side of the line is valued for its strong expressive effect.[75] The same may be said of threats and fighting words, both of which mark a thin line separating vituperative speech from speech that intentionally alarms or angers to such a degree as to cause paralyzing fear or immediate violence.[76] Because these unprotected speech categories represent a judgment that speech is punishable precisely because its expressive effects cross a line into the realm of the intolerable, it requires more than mere assertion to explain why these expressive effects also are so valuable as to cause concern whenever they are experienced unevenly throughout the speech marketplace.[77]

HARRY KALVEN, JR., A WORTHY TRADITION: FREEDOM OF SPEECH IN AMERICA 119-236 (Jamie Kalven ed., 1988); Martin H. Redish, *Advocacy of Unlawful Conduct and the First Amendment: In Defense of Clear and Present Danger*, 70 CAL. L. REV. 1159, 1166-77 (1982) [hereinafter Redish, *Advocacy of Unlawful Conduct*].

73. Brandenburg v. Ohio, 395 U.S. 444, 456 (1969) (Douglas, J., concurring).

74. *See, e.g.*, KALVEN, *supra* note 72, at 234-36 (explaining the "*Brandenburg* line").

75. *See id.*

76. *See, e.g.*, Gard, *supra* note 33, at 534, 536; Jennifer E. Rothman, *Freedom of Speech and True Threats*, 25 HARV. J.L. & PUB. POL'Y 283, 285 (2001).

77. Nor does the Court's instrumental or skewing-effect rationale square with the Court's judgments that some speech categories have so little expressive value that they are unworthy of protection and that the production of some speech is so harmful as to outweigh any expressive value that might inhere in the speech category. The former judgment manifests itself in obscenity doctrine. *See supra* notes 37-39 and accompanying text. If the notion that speech can be labeled so valueless as to be unworthy of protection is taken seriously, it logically follows that any harm caused by the uneven distribution of this speech is negligible at best. The latter judgment manifests itself in child pornography doctrine, which is grounded mainly in the notion that the overwhelming harms of production and distribution outweigh any speech value that might exist. *See supra* note 40 and accompanying text. It follows from this notion that any speech-related harms stemming from the uneven, content-based distribution of child pornography also are outweighed by overwhelming harms. As for the latter observation about child pornography, it bears noting that while child pornography doctrine is grounded partly in the assumption that such speech is unlikely to have much value and that any minimal value likely can be retained through simulation (such as by having adults play the roles of juveniles), *see* New York v. Ferber, 458 U.S. 747, 762-63 (1982), the doctrine is grounded primarily in a harm-based analysis. First, the Court has concluded that child pornography's production is overwhelmingly harmful to its juvenile participants and that such production cannot effectively be stopped unless the end product can be banned. *Id.* at 756-62. Second, the Court has concluded that the emotional damage of production is exacerbated by the juvenile participants' knowledge that the end product can be sold and consumed. *Id.* at 759.

Furthermore, while the *R.A.V.* Court's intrinsic harms rationale could be deemed to relate remotely to protected speech infringement, the Court's discussion ultimately steers clear of this rationale and thus remains inconsistent with categorization doctrine. The Court explains that "the power to proscribe [unprotected speech] on the basis of *one* content element (*e.g.*, obscenity) does not entail the power to proscribe it on the basis of *other* content elements,"[78] apparently deeming the latter intrinsically problematic. Interpreted generously, one might view this statement as evincing a subtextual concern that legislatures can harness protected content elements to unprotected speech categories and thus can treat the latter as regulatory vehicles through which to punish the former. The problem with this interpretation is that the *R.A.V.* Court clearly assumes that any speech punished through an unprotected speech regulation indeed is *unprotected* speech.[79] The Court thus relies on the notion that there simply is something intrinsically wrong with punishing unprotected speech on the basis of its content. This returns us to a fundamental dilemma: the apparent incompatibility of this notion with categorization doctrine.

Such incompatibility might best be explained through an example of the type of government activity about which the Court seems concerned: a legislature wishes to prohibit all racist speech because it detests racist ideas and wishes to convey its hostility toward racism. The legislature cannot, however, prohibit all racist speech because much racist speech is constitutionally protected. Therefore, the legislature uses an unprotected speech category as a vehicle through which it can partly satisfy its desire to squelch racist expression and demonstrate its hostility toward racist ideas.

While the *R.A.V.* Court's concern over such activity has some intuitive appeal, it crumbles rather quickly upon analysis. Government motivation to suppress particular ideas is deeply problematic in the context of protected expression precisely because such motivation manifests itself in the suppression of expression deemed valuable from any number of perspectives, including those focused on speaker or audience autonomy and those focused on the social value of speech.[80] But it is not immediately clear that these concerns are present where government partiality manifests itself in the regulation of activity deemed so worthless or so harmful as to fall outside of the First Amendment's protective zone. Indeed, the *R.A.V.*

78. R.A.V. v. City of St. Paul, 505 U.S. 377, 386 (1992).

79. *See, e.g.*, *id.* at 383-90 (describing justifications for limiting the regulation of "proscribable speech").

80. *See supra* notes 1, 14 and accompanying text (referring to intrinsic and instrumental theories of free speech value and to the reliance of the general content-distinction rule on these theories).

Court's concern with government motivation devolves, upon consideration, into a generalized concern with government partiality. Such concern sweeps so broadly as to implicate almost all areas of government regulation, as most government regulations can be characterized as intentionally disadvantaging particular ideas as they manifest themselves in regulated behaviors.[81]

One might defend the *R.A.V.* Court's intrinsic distaste for government favoritism by arguing that such favoritism infringes on free thought even if it does not infringe on free speech, because such favoritism embodies a desire to punish particular thoughts that manifest themselves in unprotected speech.[82] To the extent that the argument suggests that thought elements can never form part of a criminal or civil wrong, it runs counter to a long and generally justified practice of deeming thought elements, such as intent, part of criminal and civil wrongs.[83] The argument thus must devolve to the point that thought elements, such as viewpoints and subject matter, are not appropriate bases for defining or enhancing punishment. At this point, however, the argument intersects with an important question raised by this Article: when does a thought or speech element form part of unprotected speech, and when does it merely accompany unprotected speech and thus remain protected? To the extent that particular thought elements remain unexpressed by a speaker or otherwise do not form part of the speaker's unprotected speech, they are clearly protected. To the extent that particular thought elements do manifest themselves in unprotected speech, even contributing to the character that makes the speech unprotected (expressed racism might, for example, enhance the frightening nature of a threat),[84] they form part of the unprotected speech and would seem, by the logic of categorization doctrine, to be constitutionally punishable.

---

81. Of course, any government distinction, even within categories of unprotected speech, may be subject to an equal protection challenge under rational basis review. *See* CHEMERINSKY, *supra* note 1, § 9.2.1. But a free speech value-based explanation must be supplied if the distinction is said to raise First Amendment concerns and thus to justify a higher level of scrutiny. *Id.* § 10.1.1 (discrimination with respect to fundamental rights receives strict scrutiny).

82. Martin Redish makes a similar point with respect to motive-based regulations of prohibited conduct, such as hate crimes legislation. Martin H. Redish, *Freedom of Thought as Freedom of Expression: Hate Crime Sentencing Enhancement and First Amendment Theory*, 11 CRIM. JUST. ETHICS 29, 37-39 (1992).

83. *See, e.g.*, Frederick M. Lawrence, *Resolving the Hate Crimes/Hate Speech Paradox: Punishing Bias Crimes and Protecting Racist Speech*, 68 NOTRE DAME L. REV. 673, 717-18 (1993); Note, *Hate is Not Speech: A Constitutional Defense of Penalty Enhancement for Hate Crimes*, 106 HARV. L. REV. 1314, 1323-25 (1993).

84. *See infra* Part VI.C (invoking the hypothetical example of a regulation targeting race-based threats).

Finally, the weak and make-shift nature of the Court's intrinsic and instrumental rationales is further evidenced by the sharp inconsistency between such rationales and the Court's opinion in *Wisconsin v. Mitchell*,[85] decided in the years between *R.A.V.* and *Black*. In *Mitchell*, the Court upheld a statutory sentencing enhancement for an aggravated battery conviction based on the fact that the defendant selected his victim because of the victim's race.[86] The *Mitchell* Court relied partly on the notion that content-based regulations raise little or no constitutional problem where the regulations target "conduct unprotected by the First Amendment."[87] If neither the impact of content-based regulations nor their underlying purpose are of constitutional concern where the content element (in this case, race-based motivation) manifests itself in conduct unprotected by the First Amendment, it is difficult to see why such impact or purpose is of constitutional concern where a regulated content element manifests itself in expression deemed beyond the First Amendment's reach.

In failing to bridge the gap between its conclusions and the premises of categorization doctrine, the *R.A.V.* Court not only leaves a riddle unsolved, but also misses an important opportunity to reassess the nature and role of categorization doctrine. The Court could have explained that categorization doctrine, while necessary, is inherently risky because of the fine line separating unprotected speech content from protected, even highly valued speech content. Given this fine line, neither free speech value nor the instrumental and intrinsic concerns on which the content-distinction rests can be deemed completely absent from the realm of unprotected speech. Consideration of the relationship between protected speech and unprotected speech could also have shed light on the risk that unprotected speech categories will infringe on protected speech, whether through the chilling effect, by the harnessing of protected speech elements to unprotected speech legislation, or through the heuristic impact of linking content elements to unprotected speech categories.[88] Such analysis, too, could have provided an important link between categorization doctrine and content-distinction principles, a link missing from the Supreme Court's analysis of the issue.

III. CRITIQUE OF EXISTING SCHOLARSHIP

While some important scholarly debate has occurred in the years since *R.A.V.*, notably absent from this debate is consideration of the fine line between protected speech and unprotected speech, the impact that

85. 508 U.S. 476 (1993).
86. *Id.* at 479.
87. *Id.* at 487.
88. See *supra* notes 20-22 and accompanying text.

unprotected speech regulations might therefore have on free speech value, and the risk that unprotected speech regulations pose to protected speech. This Part summarizes and critiques the major scholarly arguments, discussing much of what is valuable in this scholarship while also discussing important problems and gaps in the scholarship.

The major scholarly arguments can be grouped into two core categories. The first category consists of those arguments that reason predominantly from existing doctrine, embracing content-distinction and/or categorization principles and considering how well *R.A.V.* fits within these principles. The second category consists of those arguments that critique *R.A.V.* for presuming the centrality of content-distinction principles in free speech doctrine in the first place, and/or that critique *R.A.V.* for overlooking the important role that equality principles should play in free speech doctrine.

A. *Arguments that Criticize or Embrace* R.A.V. *From the Perspective of Existing Content-Distinction and/or Categorization Principles*

Of those scholars who criticize *R.A.V.* from the perspective of existing content-distinction or categorization principles, a few focus largely on the point that it is illogical to apply content-distinction principles, or any level of heightened scrutiny, to speech deemed outside of the First Amendment.[89] This point, while significant, tends nonetheless to be unaccompanied by sustained analysis regarding the possible grounds on which such application could be justified.[90] Without serious confrontation of the intrinsic and instrumental rationales suggested by the *R.A.V.* Court or of other possible rationales, it is difficult to reject the argument that content-distinction principles have a role to play in the assessment of unprotected speech categories.

Another subset of scholars embraces all or part of *R.A.V.*'s doctrine and rationale from the perspective of existing content-distinction or categorization principles. Of these scholars, most fail to grapple seriously with the tensions posed between content-distinction and categorization principles, leading to an overly easy acceptance of all or part of the *R.A.V.* Court's approach. For example, Ronald Rotunda celebrates *R.A.V.* for its stance against the state's "officially suppress[ing] ideas" "[e]ven when prohibiting fighting words"[91] and for its rejection of "political correctness

89. *See, e.g.*, G. Sidney Buchanan, *The Hate Speech Case: A Pyrrhic Victory for Freedom of Speech?*, 21 HOFSTRA L. REV. 285, 296-98 (1992); David Goldberger, *Hate Crime Laws and Their Impact on the First Amendment*, 1992/1993 ANN. SURV. AM. L. 569, 573-75, 578-79 (1993); Philip Weinberg, R.A.V. *and* Mitchell*: Making Hate Crime a Trivial Pursuit*, 25 CONN. L. REV. 299, 300-09 (1993).

90. *See supra* note 89.

91. Ronald D. Rotunda, *A Brief Comment on Politically Incorrect Speech in the Wake of*

as a gloss limiting the First Amendment."[92] Rotunda's arguments largely track the *R.A.V.* Court's logic[93] without adequately explaining why the interests that underscore content-based or even viewpoint-based regulations are present when the speech suppressed is deemed so harmful or so valueless as to fall outside of the Constitution's protections. Similarly, Murray Dry embraces the *R.A.V.* Court's opinion largely by tracking the Court's emphasis on the special evils of viewpoint discrimination and by citing the dangers of "political correctness."[94] Dry's argument, too, fails seriously to confront arguments based on categorization principles. Laurence Tribe justifies *R.A.V.*'s result through analysis that is more nuanced than that of Rotunda or Dry but that similarly fails to consider the conflict between *R.A.V.* and categorization principles.[95] Specifically, Tribe explains that it is permissible to condition punishment on one's motivation to do a particular thing, such as selecting one's victim based on race, but that it is impermissible to condition punishment on one's motivation to express a particular message.[96] Tribe concludes that the St. Paul ordinance did the latter and thus that the *R.A.V.* Court's decision is justified.[97] This point fails to grapple with the argument that when only unprotected speech is targeted and impacted, any message punished is a component of speech deemed so harmful or so valueless as

---

R.A.V., 47 SMU L. REV. 9, 20-21 (1993).

92. *Id.* at 15.

93. *See id.* at 16-22. In addition to tracking the *R.A.V.* Court's logic, Rotunda states in his conclusion:

> If bad motive or the belief in bigotry may be made a crime, then the state can question defendants and subpoena records as to what newspapers or magazines they read, what books are in their library, or what they have confided to their friends. All of this information is relevant because it may show the propensity to commit the crime of bad motive, the offense of being politically incorrect. In addition, if the magazine article or book that advocates the politically incorrect viewpoint is persuasive, then it may be a crime for the author to have written it.

Id. at 21. Rotunda does not explain why laws singling out particular types of speech content necessarily require probes into other protected speech consumed or engaged in by defendants. In any event, if a law were to have this effect, its basis for invalidation would be its impact on *protected* speech, not constitutional problems inherent in the content-based regulation of *unprotected* speech. That Rotunda is partly concerned with *protected* speech infringement also is reflected in an earlier part of his essay, in which he emphasizes the narrow scope of fighting words doctrine and the fact that much "hate speech" regulation extends well beyond this scope. *Id.* at 12-15.

94. Murray Dry, *Hate Speech and the Constitution*, 11 CONST. COMMENT. 501, 506-08, 512 (1994-95).

95. *See* Laurence H. Tribe, *The Mystery of Motive, Private and Public: Some Notes Inspired by the Problems of Hate Crime and Animal Sacrifice*, 1993 SUP. CT. REV. 1, 5, 8-9, 11.

96. *Id.*

97. *Id.* at 15.

to fall outside of the Constitution's protection. From this perspective, Tribe's argument has weight only if at least one of two conditions applies: there is a convincing reason not to trust that unprotected speech categories effectively lack free speech value (or that any such value is clearly outweighed) or the St. Paul ordinance in some way punishes or threatens to punish a message encompassed in *protected* speech or thought.[98]

Scholarship embracing *R.A.V.*'s doctrine after acknowledging the tension between *R.A.V.* and categorization principles also leaves questions to be answered. An article by Edward J. Eberle recognizes, on the one hand, that there is a tension between *R.A.V.* and categorization principles,[99] that this tension threatens to dilute protected speech protections,[100] and that the *R.A.V.* Court is disingenuous in purporting to apply standard strict scrutiny rather than a significantly heightened form of strict scrutiny.[101] On the other hand, Eberle embraces the *R.A.V.* Court's doctrine on the basis that it is highly speech-protective because it reflects a strong rejection of viewpoint discrimination[102] and because the doctrine, when viewed in tandem with *Mitchell*, reflects a strong commitment to separating speech from conduct and vigorously protecting the former.[103] There are two major difficulties with Eberle's argument. First, while Eberle applauds the *R.A.V.* Court's use of heightened strict scrutiny because he deems it a strong defense against viewpoint-based regulations,[104] Eberle fails to consider seriously why viewpoint discrimination is problematic if such discrimination targets and affects only unprotected speech. While Eberle acknowledges the general tension

98. Also worth noting is Akhil Amar's argument, which breaks the *R.A.V.* opinion into two major components. One component, which Amar criticizes, is the Court's failure to take into account the special relevance of race in constitutional law. Akhil Reed Amar, *The Case of the Missing Amendments:* R.A.V. v. City of St. Paul, 106 HARV. L. REV. 124, 125-26, 130-31, 151-61 (1992). The second component, which he deems a highly positive development, is the *R.A.V.* Court's strong affirmation of certain speech-protective principles including the notions that symbolic speech is fully protected and that viewpoint discrimination is anathema to the First Amendment. *Id.* at 125, 132-46. One may interpret Amar's discussion of *R.A.V.*'s second component in one of two ways. Amar may embrace *R.A.V.*'s speech-protective principles because of their potential application to protected speech, particularly insofar as such principles were called into question in case law prior to *R.A.V. See id.* at 124-25. Alternatively, Amar may embrace the application of these principles to unprotected speech. If the latter is the case then Amar's analysis, too, fails to address the tension between *R.A.V.* and categorization principles.

99. Edward J. Eberle, *Hate Speech, Offensive Speech, and Public Discourse in America*, 29 WAKE FOREST L. REV. 1135, 1143, 1146, 1159-70 (1994).

100. *Id.* at 1146, 1159-70.

101. *Id.* at 1145, 1162-63, 1170-71; *see also infra* notes 201-09 and accompanying text (discussing the *R.A.V.* Court's disingenuous use of "heightened" strict scrutiny).

102. Eberle, *supra* note 99, at 1139, 1145, 1170-72, 1176-78.

103. *Id.* at 1169-70.

104. *Id.* at 1139, 1145, 1170-72, 1176-78.

posed between *R.A.V.* and categorization doctrine and cites the standard arguments that viewpoint discrimination is highly paternalistic and skews public discourse,[105] Eberle does not explain why the standard arguments resonate with respect to speech deemed so harmful or valueless as to be beyond constitutional protection.[106] Second, Eberle's emphasis on the importance of separating speech from conduct and of vigorously protecting only the former contradicts much of the rest of his analysis. If protections against viewpoint discrimination are so crucial that they should be invoked even in the realm of unprotected speech, it is difficult to see why this principle does not apply in the realm of conduct that expresses or is motivated by a particular viewpoint.[107]

---

105. *Id.* at 1172.

106. The closest that Eberle comes to rendering such explanation is in his contention that the *R.A.V.* Court properly separates "proscribable" speech elements from "protected" speech elements such as viewpoint or subject matter. *Id.* at 1152. Eberle suggests that this point is grounded in an intrinsic view of free speech value, whereby speech is measured by its value to the speaker, regardless of its positive or negative social impact. *Id.* at 1147, 1150. The problem with Eberle's "elements" point is that it fails to acknowledge that where "protected" elements form part of proscribable speech, such elements form a part of speech deemed so harmful or so valueless as to raise few intrinsic or instrumental concerns. If elements of proscribable speech themselves are not devoid of value, it must be because the proscribable speech of which they form a part itself is not devoid of value or even constitutes protected speech. It is this analytical link that Eberle, like the *R.A.V.* Court, does not make. Similarly, Eberle's point about free speech value runs counter to the assumption, which he does not answer, that deeming speech (or speech "elements") categorically proscribable involves a calculation that any free speech value either is absent or significantly outweighed. In any event, Eberle's free speech theory point rests somewhat on a shifting foundation. Early in his article, Eberle explains that "[t]he majority of the Court in R.A.V. preferred a nonconsequentialist view, finding that speech is valuable as an end itself, independent of any consequences that it might produce." *Id.* at 1147; *see also id.* at 1150. At another point, however, Eberle writes that

> [t]he Court's narrowing of the areas beyond protection of the First Amendment reflects the high value it places on free speech, both as an end in itself and as a means to the accomplishment of other important ends. From this vantage point, all expression is presumptively beyond governmental power. In service of this ideal, the Court seems committed to opening the channels of communication widely, allowing the transmission of more ideas and information so that we may acquire "an ever better understanding of reality" or better achieve other desirable goals, such as democratic self-government.

Id. at 1155 (quoting Michael J. Perry, *Freedom of Expression: An Essay on Theory and Doctrine*, 78 NW. U. L. REV. 1137, 1155 (1983)) (footnote call numbers omitted).

107. Eberle attempts to address such concerns by acknowledging that much nonverbal conduct is expressive in nature and arguing that such conduct can be regulated for its "conduct" elements but not for its communicative elements. *Id.* at 1198-99. However, Eberle also argues that the statute at issue in *Wisconsin v. Mitchell*, 508 U.S. 476 (1993), which enhanced aggravated battery convictions where victims were selected on the basis of their race, clearly was constitutional and distinguishable from the St. Paul ordinance at issue in *R.A.V.* Eberle, *supra* note 99, at 1197-1204.

While an article by Alan Brownstein is unusually reflective in considering whether content-distinction principles have a role to play with respect to unprotected speech, two facets of Brownstein's argument merit critique. First, Brownstein details the rationales typically offered for applying the content-distinction to protected speech, including intrinsic and instrumental concerns and fears that legislatures more readily will suppress speech when they can hone in on unpopular speech content.[108] Brownstein concludes that these rationales do not apply to speech deemed so harmful or so valueless as to be unprotected.[109] Brownstein's point usefully highlights the *R.A.V.* Court's failure to reconcile its approach with the traditional assumptions of categorization doctrine. Brownstein stops short, however, of considering whether categorization doctrine's traditional assumptions should be reassessed and whether such reassessment is the missing analytical link that eludes the *R.A.V.* Court. Brownstein thus fails to consider the possibility that there is reason to not fully trust unprotected speech categories, either because free speech value may exist within the categories or because the categories may serve as vehicles to infringe on protected speech.

Second, Brownstein's alternative rationale for closely scrutinizing content-based regulations of unprotected speech, while provocative, is highly strained. After concluding that the arguments typically raised against viewpoint-based and content-based regulations of protected speech are inapplicable to unprotected speech,[110] Brownstein explains that viewpoint-based regulations of unprotected speech should receive strict scrutiny and other content-based regulations of unprotected speech should receive intermediate scrutiny.[111] Brownstein argues that viewpoint-based regulations of unprotected speech intolerably distort the marketplace of ideas, not by directly removing valuable speech from the marketplace, but by "empowering one side of an ideological conflict with dangerous weapons that are denied to its opponents."[112] Brownstein also deems it constitutionally problematic, though to a lesser degree, to empower only those discussing selected subject matters or using other forms of selected content to abuse others through unprotected speech.[113] The core problem

---

It is difficult to understand why Eberle deems race-based motivation that manifests itself in noncommunicative content clearly punishable while deeming race-based motivation that manifests itself in unprotected speech presumptively protected.

108. Brownstein, *supra* note 13, at 589-611.

109. *Id.* at 611-24.

110. *Id.* at 589-624.

111. *Id.* at 586-87. Brownstein would apply the same logic and doctrinal standards to motive-based regulations of nonspeech conduct, such as hate crime laws. *Id.*

112. *Id.* at 628.

113. *Id.* at 633-34.

with Brownstein's argument is that its foundations simply are illogical.[114] Brownstein's "abuse" argument assumes that jurisdictions respond to clearly delineated sides in private debates and that they can, by protecting or less strictly punishing certain types of abuse, cause or intend to cause the abuse of persons clearly identified with certain sides or aspects of such debates. This argument requires major logical leaps regarding the state's ability to predict that certain content-identified speakers will be abused if others are "licensed" to use unprotected speech and equating such private abuse with state discrimination against the relevant speech content. Such leaps are simply too tenuous to support a free speech claim to the effect that states seek, through content-based regulations of threats and fighting words, to commission private actors to abuse those engaged in content-specific, protected speech or thought. Nor could Brownstein's argument withstand scrutiny if it were reframed as an argument that "unlicensed" or more heavily punished speakers are deprived of the ability to match abuse for abuse because of their expressive content. Such an argument effectively constitutes a claim of right to engage in unprotected speech—a claim that, as we have seen, has yet to be justified.[115]

Finally, the rendering of the *Black* decision provides a useful opportunity to consider a more recent take on *R.A.V.* Guy Charles offers a descriptive explanation of the *R.A.V.* Court's logic in an article that seeks to reconcile the different outcomes in *R.A.V.* and *Black* by emphasizing the role of critical race theory in the latter.[116] Charles explains that the *R.A.V.* Court's concern might have been government motivation, with the Court deeming it acceptable for the state to regulate unprotected speech for the very reasons that make it unprotected but unacceptable to regulate it for

---

114. There also are other important, but less fundamental, problems in Brownstein's arguments. First, Brownstein's logic extends only to fighting words and threats, as it would be difficult to characterize other unprotected speech categories as involving "dangerous weapons" used by one side of a conflict against the other side. *See id.* at 628. Second, Brownstein's argument seems to assume jurisdictions in which threats and fighting words generally are not punished and in which the only restrictions on such speech are content-based. Hence, Brownstein's reference to the fact that one side of a conflict is allowed, even "empower[ed]" to use "weapons" denied to the other side. *See id.* at 591. It seems far more likely that a given jurisdiction will have general threat provisions and general breach of the peace provisions encompassing fighting words, while also having more specific, content-based laws or sentence enhancements. Of course, this shortcoming does not render Brownstein's point irrelevant, as one could argue that jurisdictions with both general and content-specific provisions improperly single out specified sides of arguments for heightened punishment. The shortcoming does, however, call into question the force of Brownstein's point.

115. Frederick Lawrence and Ronald Turner raise points similar to those raised by Brownstein regarding the arguable unfairness of punishing some groups, but not others, for using unprotected speech. *See* Lawrence, *supra* note 83, at 712-13; Ronald Turner, *Hate Speech and the First Amendment: The Supreme Court's* R.A.V. *Decision*, 61 TENN. L. REV. 197, 223-24 (1993).

116. Charles, *supra* note 29, at 608-32.

unrelated content or viewpoint-based reasons.[117] Charles's point is descriptive, not normative.[118] If descriptive of the *R.A.V.* Court's rationale, however, it brings the Court no closer to a satisfying explanation for its doctrine. Where legislatures are deemed to be discriminating within the realm of "nonspeech," *Mitchell* suggests that such discrimination is subject—at most—to rational basis review.[119] If concerns justifying more rigorous scrutiny exist within the realm of unprotected speech, they remain to be explained.

B. *Arguments That Criticize* R.A.V. *for Overemphasizing the Centrality of Content-Distinction Principles and/or for Underemphasizing Equality Principles*

Steven Shiffrin criticizes the *R.A.V.* Court for deeming neutrality, and hence the content-distinction principle, the core free speech value while ignoring the proliferation of content-based judgments throughout categorization doctrine[120] and overlooking the fact that equality principles sometimes outweigh neutrality principles.[121] Shiffrin directs the bulk of his analysis toward the points that neutrality values must be considered in tandem with the values of dissent and equality in free speech jurisprudence,[122] and that such consideration yields the conclusions that much racist speech is not constitutionally protected but that it should not, for policy reasons, be fully regulated.[123] While Shiffrin emphasizes his theory of competing values in free speech jurisprudence, his foundational point regarding the reflexive nature of *R.A.V.*'s doctrine helps to further illuminate the need for probing analysis as to the propriety of categorization in free speech doctrine and as to whether, when, and why content-distinction principles have a role to play with respect to unprotected speech.

Other scholars, regardless of their views on the relative significance of neutrality as a constitutional principle, share Shiffrin's view that the *R.A.V.*

---

117. *Id.* at 590-95.

118. Similar descriptive points were helpfully suggested to me by Larry Alexander and Nelson Tebbe. *See* Schauer, *supra* note 29, at 207 ("[I]t is sounder to think . . . of a First Amendment not that protects speech, but instead that prohibits certain reasons for restricting it.").

119. *See supra* notes 85-87 and accompanying text. Indeed, any such review would be generated as a matter of equal protection law, not of free speech law, to the extent that the speech at issue is deemed nonspeech or unprotected speech. To justify a higher level of review, either First Amendment concerns or a suspect classification would have to be shown. *See supra* note 81 (discussing levels of review under equal protection law).

120. Steven H. Shiffrin, *Racist Speech, Outsider Jurisprudence, and the Meaning of America*, 80 CORNELL L. REV. 43, 56-57, 60-64, 91-93 (1994).

121. *Id.* at 59, 68-69, 91.

122. *Id.* at 84-93.

123. *Id.* at 93-102.

Court failed to take seriously the competing value of equality in the case.[124] For example, one of the arguments made by Guy Charles is that the *Black* Court applied its standards in a manner that took the harms of racism seriously while the *R.A.V.* Court applied its standards in a manner dismissive of such harms.[125] While such points are important, they take on greater clarity and significance when they are considered in tandem with the broader doctrinal framework within which courts consider race-based speech measures. A virtue of carefully limited unprotected speech categories is that they provide a framework to consider acceptable bases for speech regulation. Within this framework, courts can assess whether and when racism truly manifests itself as unprotected speech and whether, when, and why such manifestation can be treated differently than other instances of unprotected speech. Yet if such assessment is to successfully be made, one must first consider whether unprotected speech is fully unprotected and if not, what principles and standards shape its parameters.

IV. A NEW APPROACH FOR ASSESSING CONTENT-BASED REGULATIONS OF UNPROTECTED SPEECH: THEORETICAL FOUNDATIONS

This Part gives substance to the widely shared, but so far inadequately explained, instinct that there is something constitutionally troubling about content-based regulations of unprotected speech. This Part explains that this instinct only reasonably can stem from the relationship between protected speech and unprotected speech. Subpart A explains the categorization paradox. The categorization paradox embraces the ideas that categorization is necessary and desirable but that the very elements that make speech unprotected also contribute to an inescapably fine line between unprotected speech and protected, even deeply valued speech. Given the close relationship between protected speech and unprotected speech, categorization is best understood as a necessary evil in free speech doctrine. Subpart B discusses the containment principle. This principle, which follows from the categorization paradox, encompasses the notion that the scope and use of unprotected speech categories must be carefully contained in light of the potential impact of unprotected speech regulations

---

124. *See, e.g.*, Amar, *supra* note 98, at 125-26, 151-61; Charles R. Lawrence, III, *supra* note 48, at 787-804; Cass R. Sunstein, *Words, Conduct, Caste*, 60 U. CHI. L. REV. 795, 825-28 (1993).

125. Charles, *supra* note 29, at 602-03, 610-13, 629-32. For reasons explained in my discussion of *R.A.V.*'s doctrinal standards and their application, I disagree with Charles's suggestion that the *R.A.V.* Court's conclusion was indefensible. It is true, however, that the *R.A.V.* Court's language and tone suggested a dismissiveness toward claims grounded in the heightened harms of racism, even where the Court's doctrine left some opening to assess such harms—an opening which the Court took in *Black*. Certainly, the shakiness of *R.A.V.*'s theoretical foundations and of the doctrine formulated therein leaves uncertain the likely consistency and merit of any future determinations that the Court might make regarding similar questions.

on free speech value and on protected, even deeply valued speech. Subpart B details the particular risks posed by content-based regulations of unprotected speech to free speech value and to protected speech. Subparts C and D explain that the seeds for conceptualizing the categorization paradox and the containment principle can be gleaned from existing free speech theory and doctrine. Subpart E explains that social science research on evaluative heuristics bolsters an important aspect of the argument that content-based regulations of unprotected speech threaten protected speech.

A. *The Categorization Paradox*

1. The Categorization Paradox Part I: The Desirability of Categorization

The categorization paradox encompasses two major points. First, categorization is a necessary and desirable part of free speech doctrine. Second, categorization is necessarily problematic because many of the elements that justify deeming speech unprotected bear a close relationship to protected, even highly valued speech.

With respect to the first point, a limited categorization doctrine is consistent with free speech theory to the extent that such doctrine is geared toward identifying speech that not only generates overwhelming harms, but that does so in a manner generally antithetical to free speech value. Indeed, this Article earlier made the descriptive observation that categorization cases often are based on the notion—sometimes explicit and sometimes implicit—that harms caused by unprotected speech not only outweigh any free speech value but also operate in a manner antithetical to such value.[126] For example, the recent split by an en banc panel of the Ninth Circuit over the constitutional definition of "true threats" was motivated partly by competing conclusions as to when threats circumvent the preconditions of intrinsic and instrumental free speech theories by undermining threat recipients' autonomy and ability to respond to the threats.[127]

Scholars similarly have argued that particular unprotected speech categories operate in a manner antithetical to free speech value. For example, Edwin Baker, among others, has justified certain speech restrictions, such as those on threats, through an autonomy-based conception of the First Amendment that deems speech that circumvents

126. *See supra* notes 43-44 and accompanying text.

127. Planned Parenthood v. Am. Coalition of Life Activists, 290 F.3d 1058, 1085-86 (9th Cir. 2002); *id.* at 1088-89 (Reinhardt, J., dissenting); *id.* at 1098-100 (Kozinski, J., dissenting), 1106-07 (Berzon, J., dissenting).

listeners' autonomy beyond constitutional protection.[128] Scholars such as Kent Greenawalt have based similar arguments on both intrinsic and instrumental theories of free speech value, arguing that threats or other forms of speech that seek effectively to alter situations through fiat rather than persuasion are inconsistent with free speech theory and are not constitutionally protected.[129]

That speech content sometimes creates harms so overwhelming or antithetical to free speech value as to justify regulation is but a subset of the larger observation that content-based speech regulation sometimes is legitimate. These observations raise the question of what doctrinal tools courts should have at their disposal to distinguish acceptable from unacceptable content-based regulations. Because the alternative to categorization is reliance solely on one-size-fits-all balancing tests, the relatively desirable, speech-protective nature of a limited categorization doctrine is easy to discern.

Categorization doctrine is relatively speech-protective because of its conduciveness to the creation and application of fine-tuned standards. Of course, categorization is no panacea. Debates long have existed over whether and to what extent it makes a difference whether courts "categorize" or "balance."[130] And history suggests that the content of a test or standard may be less predictive of a case's result than the personal views that inform the judge's application of the relevant standards.[131] Nonetheless, tests to define specific categories of unprotected speech

---

128. C. EDWIN BAKER, HUMAN LIBERTY AND FREEDOM OF SPEECH 54–65 (1989); *see also* Christina E. Wells, *Reinvigorating Autonomy: Freedom and Responsibility in the Supreme Court's First Amendment Jurisprudence*, 32 HARV. C.R.–C.L. L. REV. 159, 161, 165-73, 179-80, 186-96 (1997).

129. *See* KENT GREENAWALT, FIGHTING WORDS 48–50 (1995); KENT GREENAWALT, SPEECH, CRIME, AND THE USES OF LANGUAGE 57–63, 67–68, 94–99 (1989).

130. *See, e.g.*, T. Alexander Aleinikoff, *Constitutional Law in the Age of Balancing*, 96 YALE L.J. 943, 943-44, 960-63, 966-67 (1987); John Hart Ely, *Flag Desecration: A Case Study in the Roles of Categorization and Balancing in First Amendment Analysis*, 88 HARV. L. REV. 1482, 1500-03, 1506-08 (1975); Robert F. Nagel, *Liberals and Balancing*, 63 U. COLO. L. REV. 319, 319-24 (1992); Kathleen M. Sullivan, *Post-Liberal Judging: The Roles of Categorization and Balancing*, 63 U. COLO. L. REV. 293, 293-96, 306-09 (1992).

131. *Compare* Schenck v. United States, 249 U.S. 47 (1919) (majority opinion, applying clear and present danger test to uphold a conviction, by Justice Holmes), *with* Abrams v. United States, 250 U.S. 616, 624-31 (1919) (dissenting opinion, arguing that conviction should be overturned under clear and present danger test, by Justice Holmes). Certainly, Justice Holmes's differing positions in the two cases could be attributed to differences in the facts. But one also could conclude that Justice Holmes grew warier of government control of speech in the eight months between the two opinions and that he therefore applied the same test more leniently in *Abrams* than in *Schenck*. *See, e.g.*, Martin H. Redish, *Advocacy of Unlawful Conduct*, *supra* note 72, at 1167-68 ("Commentators sense that Holmes dramatically shifted his emphasis with his dissent in *Abrams v. United States*.").

logically seem more conducive to identifying narrow conditions on which a lack of constitutional protection hinges than do generally applicable balancing tests.[132] History seems to bear this perception out. For example, while the story of incitement doctrine's development at its core is a story of a deepening theoretical commitment to open political discourse, such commitment manifests itself doctrinally in a set of specific standards defining "inciting" speech.[133] Similarly, while the definition and propriety of "fighting words" remain cloudy, there is widespread agreement that the term's meaning has narrowed considerably over the years, paralleling a deepened theoretical commitment on the part of the Supreme Court to unfettered social discourse.[134]

Finally, a limited categorization doctrine serves as a partial defense against the dilution of generally applicable doctrinal standards. While categorization doctrine is vulnerable to overly speech-restrictive implementation, the doctrine has the virtue of forcing open judicial discussion regarding clearly momentous decisions to deem particular speech categories unprotected. The alternative, providing the judiciary with a single set of one-size-fits-all standards, risks the more insidious judicial disfavoring of particular speech forms through unusually harsh applications of standard doctrinal tests. Such an alternative also risks diluting general speech protections as the Court strains to interpret existing standards in a way that enables it to uphold regulations of disfavored, but technically protected speech.[135]

2. The Categorization Paradox Part II: Categorization as a Necessary Evil

The same factors that make categorization necessary also make it a risky and necessarily imperfect endeavor. Paradoxically, the very

---

132. *See, e.g.*, Ely, *supra* note 130, at 1500 & n.75; Cass R. Sunstein, *Government Control of Information*, 74 CAL. L. REV. 889, 909 (1986).

133. *See* Brandenburg v. Ohio, 395 U.S. 444, 447 (1969). *But see* Redish, *Advocacy of Unlawful Conduct*, *supra* note 72, at 1175-82 (critiquing ambiguities in the *Brandenburg* test and offering an alternative formulation).

134. *See, e.g.*, Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows*, *supra* note 1, at 1347 n.36 (citing Texas v. Johnson, 491 U.S. 397, 409 (1989); Chaplinsky v. New Hampshire, 315 U.S. 568, 572 (1942); Gard, *supra* note 33, at 534.

135. *See* R.A.V. v. City of St. Paul, 505 U.S. 377, 403 (1992) (White, J., concurring in the judgment); *id.* at 415 (Blackmun, J., concurring in the judgment); Sunstein, *supra* note 132, at 909; *cf.* Vincent Blasi, *The Pathological Perspective and the First Amendment*, 85 COLUM. L. REV. 449, 476-80 (1985) (advocating that free speech doctrine be viewed from a "pathological perspective," or from the perspective of considering how best to protect free speech in times of crisis when pressures to censor are particularly strong. From this perspective, Blasi observes "a close correlation between the ambit of coverage and the ability of courts to keep doctrine simple, informed by tradition, and dominated by principles.").

characteristics that can push speech into the realm of the unprotected are characteristics that might, at a different level of intensity or in a slightly altered context, be highly valued from the perspective of free speech theory. Consider the example of incitement doctrine. As noted earlier, incitement doctrine is so complicated precisely because speech that intentionally incites imminent illegal activity may involve impassioned calls to action grounded in social or political views.[136] Where such expression is so closely "brigaded with action"[137] that it intentionally and imminently incites illegal activity, it crosses a fine line separating punishable triggers to action from protected, even deeply valued expression that passionately advocates illegal activity more abstractly.[138] Yet the crossing of this line should not obscure the close relationship between "incitement" and protected, often deeply valued speech. Similarly, this Article has earlier referred to a recent split within the Ninth Circuit, en banc, as to the definition of a punishable threat.[139] The disagreement between the judges in this case was so heated precisely because the dissenting judges insisted that some speech very close to the "threats" line—including the speech at issue there—not only falls short of the line but in fact is highly valued for its passionate, vehement nature.[140] And one need look no further than the trail of obscenity prosecutions against the late comedian Lenny Bruce to understand that the same qualities that make speech valueless filth to one observer can make speech a highly valued rejection of a culture's sacred cows to another observer.[141]

The categorization paradox thus stems from the fact that categorization, while a necessary and even speech-protective alternative to one-size-fits-all balancing tests, is necessarily imperfect because of the inescapably fine line between unprotected speech and protected, even deeply valued speech. Misjudgments as to how to draw and police this line are inevitable, as reflected in the shifting of the line over the years. Similarly, it is disingenuous to suggest that free speech value truly ceases to exist in any meaningful way the minute that speech is deemed unprotected. The line between unprotected and protected, even highly valued speech is too fine and too murky to draw that conclusion.

136. *See supra* note 72 and accompanying text.

137. *See Brandenburg*, 395 U.S. at 456 (Douglas, J., concurring).

138. *See, e.g.*, KALVEN, *supra* note 72, at 234-36 (explaining the "*Brandenburg* line").

139. *See supra* notes 44, 126 and accompanying text.

140. *See supra* notes 44, 126 and accompanying text.

141. *See generally* RONALD K. L. COLLINS & DAVID M. SKOVER, THE TRIALS OF LENNY BRUCE (2002) (describing the rise and fall of the iconic American comedian).

B. *The Containment Principle and the Risks Posed by Content-Based Regulations of Unprotected Speech*

The categorization paradox has two consequences. On the one hand, it is important that courts maintain categorization as a tool for discerning those rare categories of speech that can be deemed unprotected. If this tool is to mean anything, it is important that legislatures have substantial leeway to regulate within these categories. Such categories can be legislated against in their entirety. In addition, reasonable leeway should exist for legislatures to legislate within subcategories of these categories. On the other hand, the categorization paradox tells us that we have to expect misjudgments, even abuse in the definition and regulation of these speech categories. Misjudgments by courts are inevitable, both as to which speech categories should be unprotected and also as to whether and when speech falls within such categories. Abuse by legislatures and law enforcement personnel who use such categories to chill or prosecute protected speech also are inevitable. Thus, some limits on the use of unprotected speech categories are called for. The limits should be geared toward minimizing the impact of categorization doctrine on free speech value in two sets of cases: first, in cases where speech is unprotected but has significant value and thus reflects a flaw with the unprotected speech category itself; and second, in cases where categorization serves as a vehicle to infringe on protected speech.

The content-distinction doctrine is an important tool for containing these two types of impact. This is because content-based regulations of unprotected speech are uniquely poised to create such impact. This Subpart explains how content-based regulations of unprotected speech threaten to impact free speech value and protected speech.

First, because free speech value cannot reliably be deemed absent from the realm of unprotected speech, the free speech value-based concerns that underscore the content-distinction rule are not fully absent from the realm of unprotected speech.[142] This insight provides the missing link that eludes the *R.A.V.* Court. The *R.A.V.* Court begins its analysis by reciting and embracing the standard justifications for categorization doctrine, explaining that unprotected speech categories long have been deemed to have "'such slight social value as [steps] to truth that any benefit that may be derived from them is clearly outweighed.'"[143] After this confident

142. For detailed discussions regarding the concerns underlying the content-distinction in the realm of protected speech, see sources cited *supra* note 13. *But see generally* Martin H. Redish, *The Content-Distinction in First Amendment Analysis*, 34 STAN. L. REV. 113 (1981) (criticizing the content-distinction as unjustifiable even in the realm of protected speech).

143. R.A.V. v. City of St. Paul, 505 U.S. 377, 383 (1992) (quoting Chaplinsky v. New Hampshire, 315 U.S. 568, 572 (1942)).

assertion, however, the Court leaps to the conclusion that unprotected speech categories are not, in fact, "invisible to the Constitution."[144] As explained earlier, the Court's conclusion is illogical if the Court continues to make the assumption, which it purports to make, that free speech value reliably can be deemed absent or overwhelmingly outweighed in the realm of unprotected speech. The Court's conclusion makes sense, however, if one begins with the premise that unprotected speech categories are a necessary evil—that they are important and valid tools of free speech doctrine, but that there are reasons not fully to trust the judgments that they embody. From this perspective, one can proceed reasonably to the conclusion that the use of categorization must be contained and that content-distinction principles have some role to play in the realm of unprotected speech. Specifically, one can understand, from this perspective, the *R.A.V.* Court's invocation of the standard, free speech value-based concerns underscoring the content-distinction: intrinsic concerns with content-based regulation and instrumental concerns about the skewing effect that such regulation has on the marketplace of ideas.[145] One can also understand the relevance of a third danger: the risk that legislatures more readily will suppress speech when they can hone in solely on unpopular speech content.[146] Taken together, these concerns suggest the following risk posed by content-based regulations of unprotected speech: Such regulations make it too easy for legislatures to hone in on particularly unpopular unprotected speech, thus skewing public discourse, causing intrinsic harms to speakers and causing intrinsic social harms borne of legislative bad faith.

Of course, the containment principle takes us well beyond the insight that content-distinction principles have a role to play in the realm of unprotected speech. The bolder insight to which the containment principle lends itself is the insight that unprotected speech categories have an impact on their close cousin: protected speech. In particular, content-based regulations of unprotected speech pose three major risks to protected speech.

The first such risk stems from the long-acknowledged chilling effect that unprotected speech has on protected speech. As discussed further in Subpart D, existing doctrine manifests a limited understanding of the relationship between protected speech and unprotected speech, although this understanding rarely impacts inquiries beyond those concerning the definitions of unprotected speech categories. Nonetheless, this understanding gives rise to widespread acknowledgement of the chilling effect, or the risk that unprotected speech categories will inhibit more

144. *Id.*

145. *See id.* at 383-90; *see also* sources cited *supra* note 13.

146. *See* Brownstein, *supra* note 13, at 608-10.

speech than they technically prohibit because speakers fearful of punishment will err on the side of caution.[147] This risk exacerbates the problems posed by content-based regulations of unprotected speech. To the extent that content-based regulations of unprotected speech may chill protected speech, any such chilling effect will be distributed in an uneven, content-based manner. Because such an effect operates within the realm of protected speech, the standard, free speech value-based concerns underscoring the content-distinction rule are raised.[148]

The second risk to protected speech is that content-based regulations of unprotected speech will use unprotected speech as a vehicle to punish persons for protected speech that accompanies but that does not form part of the unprotected speech. This danger is reflected in the following hypothetical example drawn by the *R.A.V.* Court itself:[149] a statute regulating "threats against the President that mention his policy on aid to inner cities."[150] Were a speaker to say "I am going to kill the President if he does not alter his policy on aid to inner cities," there is a strong argument to the effect that every element of the statement is part of a threat. The reference to inner-city aid policy constitutes the very contingency of the threat. On the other hand, imagine the following hypothetical rant by a caller to a radio talk show:

> I hate the President. Somebody should kill him. Hell, *I'll* kill him. That's right. Tell his secret service guards to watch their man's back, because I'm going to get him. It's about time that somebody did. I mean, look at our international policy. Look at our domestic policy. My god, look at the President's policy on aid to inner cities. That's right, somebody's going to get him soon. And that somebody will be me.

While this rant likely includes a punishable threat,[151] it is far less clear that the rant's reference to inner city aid policy is part of the threat. To the contrary, the reference seems part of a protected, verbal backdrop of anger

147. *See supra* note 20.

148. *See supra* note 13 and accompanying text (citing sources discussing the standard justifications for the content-distinction rule).

149. The *R.A.V.* Court interprets this hypothetical statute as regulating only unprotected speech but as doing so in a constitutionally problematic manner. *See R.A.V.*, 505 U.S. at 388. In actuality, the hypothetical statute should be understood to pose a risk to protected speech, as explained herein.

150. *Id.* at 388.

151. *Cf.* United States v. Kelner, 534 F.2d 1020, 1020-22, 1025, 1028 (2d Cir. 1976) (upholding a conviction on the ground that a televised statement promising to kill Yassir Arafat while he visited the United States was a "true threat"). *But see* Planned Parenthood v. Am. Coalition of Life Activists, 290 F.3d 1058, 1097-99 (9th Cir. 2002) (Kozinski, J., dissenting) (arguing that speech that is publicly addressed and political in nature is very unlikely to constitute a "true threat"); *id.* at 1091-92, 1094-96 (Beezon, J., dissenting) (same).

and heated rhetoric, accompanying but not constituting part of the threat. This example demonstrates the risk that legislators will use content-based regulations of unprotected speech to punish protected speech by heightening punishment for (or conditioning punishment of unprotected speech on) protected speech that is likely to accompany, but not to form part of, unprotected speech.[152]

A similar danger is reflected in the "freedom of thought" argument mentioned earlier.[153] As noted earlier, it is not necessarily problematic for legislatures to target thought as it manifests itself in unprotected speech.[154] For example, to the extent that racism is clearly manifest in the conveyance of a threat and clearly heightens the threatening nature of the speech, it likely can be punished. On the other hand, to the extent that one's use of unprotected speech triggers the punishment of unexpressed racist thoughts or expressed racist thoughts unrelated to the unprotected speech at issue, then the racism or other content element at issue serves as a vehicle to punish protected speech or thought.

The third major risk to protected speech is that content-based regulations of unprotected speech will serve as heuristic devices, hinting to speakers and factfinders that the thin line between protected and unprotected speech is crossed where a targeted content element is present. The statute at issue in *Virginia v. Black*,[155] a case discussed in more detail in Part V, exemplifies this risk. The *Black* statute prohibited threats conveyed through cross-burning.[156] Given the relative salience of the cross-burning factor and the relative murkiness of identifying and applying the elements of a protected threat,[157] such a statute bears a risk that the

---

152. Other examples abound, particularly given the likelihood that heated political or social rhetoric will accompany some speech that could be classified as threatening speech, inciting speech or fighting words. *See, e.g.*, *supra* note 151 (citing *Kelner* and *Planned Parenthood*). While *Kelner* involved general threats statutes, *see* 534 F.2d at 1020, it calls to mind the possibility of a hypothetical statute punishing "threats to a foreign politician that mention the politician's position on Palestinian statehood." *Planned Parenthood* involved the subject-specific statute of the Freedom of Access to Clinic Entrances Act (FACE) regulating threats of force against persons based on their provision of reproductive health care services. *Planned Parenthood*, 290 F.3d at 1062. *Planned Parenthood* did not, however, involve an *R.A.V.*-style challenge to FACE's subject matter specificity. Such challenges to FACE are discussed in Brownstein, *supra* note 13, at 558-84, 636-38.

153. *See supra* note 82 and accompanying text.

154. *See supra* notes 82-84 and accompanying text.

155. 538 U.S. 343 (2003).

156. *Id.* at 347-48.

157. *See, e.g.*, Watts v. United States, 394 U.S. 705, 707-08 (1969) (discussing the debate over the meaning of "true threats"); *Planned Parenthood*, 290 F.3d at 1086; *id.* at 1088-89 (Reinhardt, J., dissenting) (arguing over the definition of a "true threat"); *id.* at 1099-1100 (Kozinski, J., dissenting) (same); *id.* at 1107 (Berzon, J., dissenting) (same); Rothman, *supra* note 76, at 286-89 (discussing the debate among courts over the definition of "true threats").

cross-burning factor will create a perception on the part of speakers, factfinders and law enforcement agents that cross-burning is prohibited per se or that it is so likely to be punished that it categorically must not be engaged in. Indeed, in the facts underlying *Black*, a sheriff told Black that he was being arrested because "'there's a law in the State of Virginia that you cannot burn a cross.'"[158] While this statement could have been motivated by another element of the Virginia statute making cross-burning prima facie evidence of a threat, it seems likely that the mere existence of a "threatening cross-burning" statute also could give rise to such statements and arrests. Indeed, Justice Souter, partly dissenting in *Black*, lent highly refreshing insight to the case with a statement to the effect that the Virginia statute might constitute a "subtle effort to ban . . . even . . . nonthreatening cross burning."[159] Justice Souter's point was tied partly to his analysis of the statute's prima facie evidence provision.[160] Nonetheless, the notion that the judgment of speakers, factfinders, and law enforcement officers could be skewed toward chilling, conviction, or prosecution by a statutory cross-burning factor or a similar content-based provision is entirely plausible, regardless of whether such provision is accompanied by a prima facie evidence provision.

C. *Theoretical Support for the Categorization Paradox, for the Containment Principle, and for Conceptualizing the Risks Posed to Protected Speech*

Underlying the categorization paradox and the containment principle is a fundamental understanding of human fallibility and a consequent distrust of government speech regulation. Such distrust is consistent with the basic suspicion of government competence and trustworthiness in the realm of speech regulation that runs through all major theories of free speech value. In the realm of intrinsic free speech theory, for example, Martin Redish's claim that the only true free speech value is self-realization[161] logically must rest in part on an assumption about the untrustworthiness of government speech regulation. This is exemplified by Redish's emphasis on the importance of free speech for facilitating listeners' self-realization. Responding to Edwin Baker's argument that commercial speech constitutionally can be regulated because it stems from a profit motive and does not further speakers' self-realization, Redish explains that this argument "fail[s] to acknowledge that individuals may develop their personal and intellectual faculties by receiving, as well as by

158. *Black*, 538 U.S. at 349.

159. *Id.* at 384 (Souter, J., concurring in the judgment in part and dissenting in part).

160. *Id.* at 385-87 (Souter, J., concurring in the judgment in part and dissenting in part).

161. Redish, *Value of Free Speech*, *supra* note 1, at 594.

expressing."[162] While Redish's argument responds effectively to the point that some speech does not further speakers' self-realization, it does not respond effectively to the point that some speech poses a conflict between the self-realization of speakers and that of listeners. Racist speech, for example, might further speakers' self-realization at the same time as it stunts listeners' self-realization by making some feel fearful, unwelcome in their community, or disempowered from speaking freely. Redish's argument is much more forceful, however, if it rests in part on a distrust of government's ability to decide which speech will and will not further self-realization. From such an assumption, it follows that a system of free speech, while imperfect, is the best and safest route for facilitating the self-realization of all.[163] Indeed, in a subsequent article, Redish and Howard Wasserman argue that government distrust or "negative" speech theory complements such "positive" theories as self-realization theory, with "negative theories act[ing] as a bodyguard for the positive theories."[164]

Similarly, the marketplace of ideas theory, which epitomizes an instrumental approach to free speech theory, rests largely on the premises that "persons are fallible and thus should never be trusted formally to declare truth and to close off all debate on a matter," and that the "tendency to abuse power in the realm of free speech is quite natural."[165] While marketplace of ideas theory also rests on optimistic assumptions about the ability of persons to reason and to "discover truth" once speech

---

162. *Id.* at 620.

163. Redish's self-realization theory also can be understood to rely indirectly on distrust of government insofar as Redish grounds his theory in basic democratic principles. Specifically, Redish argues that democracy is designed to foster individual self-realization rather than to foster democracy for its own sake, and that democratic principles thus demand that individuals be left free to disseminate and consume speech on all matters impacting their development. *Id.* at 601-11. To the extent that concepts of democracy and self-government themselves rest in large part on a distrust of concentrated government power, so too is Redish's theory grounded partly in this distrust. *See* Heidi Kitrosser, *Secrecy in the Immigration Courts and Beyond: Considering the Right to Know in the Administrative State*, 39 HARV. C.R.–C.L. L. REV. 95, 128 n.183 (2004) [hereinafter Kitrosser, *Secrecy in the Immigration Courts*].

164. Martin H. Redish & Howard M. Wasserman, *What's Good for General Motors: Corporate Speech and the Theory of Free Expression*, 66 GEO. WASH. L. REV. 235, 244 (1998). Redish and Wasserman continue: "In other words, the purpose of the negative theories is to assure that the positive benefits fostered by the commitment to free expression can be attained, free from undermining by hostile external forces." *Id.* Of course, one might argue that negative theory does a greater than even share of the heavy lifting in its relationship with positive theory, given the importance of the government distrust assumption to the conclusion that positive theory's benefits are more likely to be realized by letting speech flow freely rather than by letting government decide when such benefits will be realized.

165. Kitrosser, *Secrecy in the Immigration Courts*, *supra* note 163, at 128 (citing JOHN STUART MILL, *On Liberty*, *in* THE PHILOSOPHY OF JOHN STUART MILL 185, 205, 208-18 (Marshall Cohen ed., 1961) (1859)).

is permitted to flow freely,[166] these assumptions are accompanied by cautionary assumptions about government's desire to stunt the free flow of speech and the incompetence of government to base this stunting on reliable conceptions of "truth" or "falsehood."[167]

The work of Frederick Schauer highlights the reliance of traditional free speech theories on distrust of government's ability to draw nonabusive distinctions between speech that should be suppressed and speech that should be protected. Schauer argues that negative justifications, or justifications based on distrust of government decisionmaking, not only are common components of major instrumental and intrinsic theories of free speech value, but also constitute the most sensible part of these theories.[168] Schauer concludes that "the most persuasive argument for a Free Speech Principle is what may be characterized as the argument from governmental incompetence."[169]

Finally, viewing critical race and gender-based theories of speech suppression through the lens of government distrust illuminates a major contradiction in such theories while also illuminating how such a contradiction may be reconciled with the theories' useful insights. Most of the critical race and gender literature that advocates suppressing racist or misogynistic speech is grounded in a belief in a deeply racist and sexist social, legal, and political structure.[170] On the one hand, such grounding is

---

166. *Id.*

167. *Id.*; SCHAUER, *supra* note 43, at 33-34. More modern theories of free speech, too, generally are grounded in some optimistic assumptions about human nature coupled with pessimistic assumptions about the reliability of government to regulate speech responsibly. Lee Bollinger's theory that free speech cultivates a characteristic of tolerance among social actors rests on the twin premises that persons have a capacity for profound intolerance that can manifest itself through majority rule, and that speech represents an area in which enforced tolerance can maximize human tendencies toward reasoned understanding. LEE C. BOLLINGER, THE TOLERANT SOCIETY 107-44 (1986). Another example is found in Steven Shiffrin's theory that the First Amendment should be understood in large part as a mechanism to celebrate, cultivate, and protect a spirit of dissent and nonconformity among persons. STEVEN H. SHIFFRIN, THE FIRST AMENDMENT, DEMOCRACY, AND ROMANCE 5-8, 86–109, 140-69 (1990). Shiffrin's theory is grounded in the twin premises that there is a strong human tendency toward conformity and toward the enforcement of imperfectly realized or imperfectly understood "truths," *see id.* at 92-95, and that a freedom of dissent helps stem tendencies toward totalitarianism and intellectual stagnation through the embrace of both reasoned deliberation and a romantic spirit of nonconformity. *See id.* at 91-96, 142, 159-61.

168. SCHAUER, *supra* note 43, at 33-34, 45, 70-72.

169. *Id.* at 86.

170. *See* Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows*, *supra* note 1, at 1383 & nn.232-33 (citing RICHARD DELGADO & JEAN STEFANCIC, MUST WE DEFEND NAZIS? HATE SPEECH, PORNOGRAPHY, AND THE NEW FIRST AMENDMENT 43, 83, 85, 86 (1997); CATHARINE A. MACKINNON, *Francis Biddle's Sister: Pornography, Civil Rights, and Speech*, *in* FEMINISM UNMODIFIED 146-49, 154-55, 174, 178 (1987); Charles R. Lawrence, III, *If He Hollers Let Him Go: Regulating Racist Speech on Campus*, 1990 DUKE L.J. 431, 461; Mari J. Matsuda, *Public Response to Racist Speech: Considering the Victim's Story*, *in* WORDS THAT WOUND: CRITICAL RACE

suggestive of the profound harms that such speech can cause. On the other hand, such grounding suggests that the government is particularly likely to be inept or abusive in purporting to regulate racist or sexist speech or thought, as the government itself is both product and perpetuator of the social framework identified by critical theorists.[171] The insights of critical theorists thus do not undermine theoretical support for the categorization paradox, the containment principle, or fears that government will abuse the power to regulate speech. Rather, such insights significantly bolster these points. Yet the containment principle, when considered in tandem with categorization doctrine, suggests a way that free speech doctrine can embrace critical race and gender theories' useful insights while rejecting their contradictions. Specifically, free speech doctrine can deem content-based regulations of unprotected speech legitimate outlets to punish the harms of racism or sexism to the extent that such racism or sexism genuinely manifests itself in unprotected speech and heightens the harms thereof. Indeed, the insights of critical theory can help shed light on when and how racism or sexism heightens the harms of unprotected speech.[172] At the same time, free speech doctrine can respond to the dangers implicit in such regulations by filtering them through standards designed to ensure that they do not infringe on protected speech or otherwise unduly impact free speech value.

D. *Doctrinal Support for the Categorization Paradox, for the Containment Principle, and for Conceptualizing the Risks to Protected Speech*

Seeds of support for the categorization paradox, the containment principle, and the risk that legislatures will harness unprotected speech categories to encroach on protected speech also can be derived from existing doctrine. Perhaps the most obvious source of such support is the Supreme Court's emphasis on the risk of a chilling effect from defamation law.[173] In *New York Times Co. v. Sullivan*,[174] the Court famously held that false criticism of public officials' official conduct is not punishable as defamation unless the claimant can prove that "the statement was made with 'actual malice'—that is, with knowledge that it was false or with reckless disregard of whether it was false or not."[175] The *New York Times* Court stressed the fine line between highly valued, vigorous criticism of

---

THEORY, ASSAULTIVE SPEECH, AND THE FIRST AMENDMENT 20, 23-24 (Robert W. Gordon & Margaret Jane Radin eds., 1993)).

171. *See* Kitrosser, *supra* note 1, at 1384-85, 1385 n.237.

172. *See* Charles, *supra* note 29, at 628-32.

173. Many thanks to Marty Redish for suggesting the example of defamation law.

174. 376 U.S. 254 (1964).

175. *Id.* at 279-80.

public officials and false, defamatory statements about such officials.[176] Given this fine line, the Court noted that a less stringent rule would intolerably chill nondefamatory speech, as

> would-be critics of official conduct may be deterred from voicing their criticism, even though it is believed to be true and even though it is in fact true, because of doubt whether it can be proved in court or fear of the expense of having to do so. They tend to make only statements which "steer far wider of the unlawful zone."[177]

A stringent rule is necessary, in short, to give protected speech the "'breathing space'" necessary to survive.[178]

The Supreme Court also has coupled its observation that unprotected speech must be defined strictly to prevent encroachment on protected speech with the observation that even expressions of official disapproval of speech can chill the dissemination and consumption of such speech. These combined concerns form the basis of the decision in *Bantam Books, Inc. v. Sullivan*,[179] in which the Court deemed unconstitutional the practices of Rhode Island's "Commission to Encourage Morality in Youth."[180] Such practices included the Commission notifying book distributors when the Commission deemed certain books or magazines "objectionable for sale, distribution or display" to minors.[181] Although refusal to cooperate with the Commission technically violated no law, the Court found that the Commission "deliberately set about to achieve the suppression of publications deemed 'objectionable' and succeeded in its aim" through various means, including by following written requests for cooperation with visits from law enforcement officers inquiring as to the steps that distributors had taken toward cooperation.[182] Noting that the Commission did not direct its pressure solely toward obscene works,[183] the Court stressed the fine line between protected and unprotected speech and the need to police that line vigorously.[184] Citing free speech's vulnerability to "gravely damaging yet barely visible encroachments," the Court explained that its "insistence that regulations of obscenity scrupulously embody the most rigorous procedural safeguards . . . is . . . but a special

176. *Id.* at 271-73.
177. *Id.* at 279 (quoting Speiser v. Randall, 357 U.S. 513, 526 (1958)).
178. *Id.* at 271-72 (quoting NAACP v. Button, 371 U.S. 415, 433 (1963)).
179. 372 U.S. 58 (1963).
180. *Id.* at 59.
181. *Id.* at 61.
182. *Id.* at 62-64, 66-68.
183. *Id.* at 64.
184. *Id.* at 65-66.

instance of the larger principle that freedoms of expression must be ringed about with adequate bulwarks."[185] The Court went on to explain that even non-binding suggestions of official disapproval can chill protected speech.[186]

Of course, there are important differences between the facts of *New York Times*, *Bantam Books*, and cases such as *R.A.V.* and *Black*. And the Court has refused to extend the *Bantam Books* principle itself very far, holding, twenty-six years after *Bantam Books*, that the federal government's practice of labeling selected foreign films as "propaganda" and requiring such films to fulfill certain registration, filing, and disclosure requirements does not unconstitutionally restrain speech.[187] Nonetheless, *Bantam Books* and *New York Times* reflect doctrinal seeds of support for the categorization paradox and the containment principle. Both cases reflect a general wariness by the Court of categorization doctrine's power and a resulting recognition by the Court that unprotected speech categories must be carefully confined, lest they encroach on protected speech. Furthermore, *Bantam Books* reflects a concern that official actions that do not directly constrain protected speech—but that hint at the illegality of such speech—can encroach on protected speech in a manner more dangerous than direct regulation.[188]

E. *Support in Social Science for Recognizing a Risk Posed by Content-Based Regulations of Unprotected Speech to Protected Speech*

Social science research on evaluative heuristics bolsters an important aspect of the argument that content-based regulations of unprotected speech threaten protected speech. Specifically, such research bolsters the argument that such regulations can serve as heuristic devices, suggesting to speakers, factfinders, and others that speech merely close to the unprotected line crosses the line where a targeted content element is present. The point is not that social science support is a necessary component of a free speech argument or that the insights of social science trump the assumptions of free speech theory.[189] The point, rather, is that the insights of social science help to confirm and explain the common

185. *Id.* at 66.

186. *Id.* at 66-70.

187. *See* Meese v. Keene, 481 U.S. 465, 484-85 (1987).

188. *See Bantam Books*, 372 U.S. at 69-70.

189. Indeed, one might argue that the assumptions about human nature made in free speech theory are not always intended to be descriptively accurate so much as they are intended to reflect an ideal that must be assumed for other reasons, such as a commitment to human dignity or a fear of the oppressive and paternalistic government that might result from different assumptions. *Cf. supra* Part IV.C (arguing that a fear of government abuse and incompetence underscores all major theories of free speech value).

sense intuition, already consistent with free speech theory and doctrine, that content-based regulations of unprotected speech can serve as heuristic devices that threaten protected speech.[190]

Social science research on evaluative heuristics suggests that individuals subconsciously adopt decision making shortcuts to compensate for limited time, knowledge, and analytical skills.[191] While it is not inherently illogical to rely on proxies to identify factors that generally correlate with such proxies, individuals may choose their proxies poorly and thus may make mistakes in assuming that certain factors indicate the existence of other factors. Heuristics literature identifies common mistakes that individuals make in treating the existence of certain factors as proxies for the existence of other factors.

The heuristic most relevant to this Article's argument is the "anchoring heuristic," whereby individuals tend to tailor their answers to questions so that the answers bear a close relationship to facts that they are given, even where it is clear that the facts have no relationship to the questions. Paul Horwitz explains a classic anchoring experiment and basic conclusions drawn from the experiment:

> Consider this experiment, conducted by Tversky and Kahneman. A set of subjects was confronted with a wheel of fortune, which was spun and landed on sixty-five. They were first asked whether the percentage of African countries in the U.N. was more or less than sixty-five percent, and then what the exact percentage was. These respondents gave a median estimate of forty-five percent. Another set of subjects was faced with a wheel of fortune that landed on ten. They gave a median estimate of twenty-five percent. This example illustrates the phenomenon of anchoring and adjustment: individuals who are asked to estimate a number often "anchor" on an original starting value, and then fail to adjust sufficiently up or down from this original starting point. This is true even when the anchor is arbitrarily chosen, or is outrageously extreme. Juries deliberating on damage awards may center on an "anchor" figure, thus skewing the damage determination toward that number, even if it is too low or high.[192]

---

190. *Cf.* Paul Horwitz, *Free Speech as Risk Analysis: Heuristics, Biases, and Institutions in the First Amendment*, 76 TEMP. L. REV. 1, 8 (2003) (by viewing free speech law through the lens of behavioral science, we can understand how the law reflects recognition of "predictable shortcomings in our ability to perform accurate risk analysis").

191. *Id.* at 12; Jeffrey J. Rachlinski, *The Uncertain Psychological Case for Paternalism*, 97 NW. U. L. REV. 1165, 1165-66 (2003).

192. Horwitz, *supra* note 190, at 18-19.

Similarly, as Bill Eskridge and John Ferejohn explain, anchoring research suggests that factfinders may "latch[] onto the most emotionally salient factor [in a case] as an anchor, which [they then] adjust[] by other factors, rather than [deem] trumped by these factors."[193]

Anchoring research bolsters this Article's argument that content-based regulations of unprotected speech may serve as heuristic devices, suggesting to speakers, factfinders, and law enforcement officials that speech merely close to the unprotected line crosses the line where a targeted content element is present. Consider again the example of threats conveyed through cross-burning.[194] Determining whether something constitutes a threat is a murky and complicated task, particularly given the often fine line between unprotected speech and protected, even deeply valued, speech.[195] Determining whether someone has burned a cross, however, is a fairly straightforward matter. Given the relative simplicity and emotional salience of the latter determination, there is a risk that factfinders will anchor their assessment of the former in an affirmative determination of the latter, at least in close cases. Furthermore, there is a greater risk that heuristic skewing will occur outside of the factfinding context, given the absence of procedural safeguards and intensive deliberation outside of that context. For example, as the facts underlying *Black* suggest, law enforcement officials and speakers may focus on the most salient aspect of a statute concerning threats conveyed through cross-burning to conclude that cross-burning generally is prohibited.[196]

Of course, the existence of cross-burning or other content elements could prove emotionally salient for factfinders even in cases prosecuted under general statutes. For example, factfinders may punish cross-burners for the ideological message conveyed by cross-burning by deeming their speech threatening under a general threats statute. This type of risk is not fully avoidable, although it can be reduced through judicial protections such as jury instructions and standards for evidentiary admissibility. Content-based statutes, however, add additional risks to the mix. Such statutes threaten to serve as legislative hints to the effect that a targeted content element's presence is a proxy for unprotected speech. Such statutes also give factfinders, speakers, and law enforcement officials statutory elements on which to anchor ideological biases or fears of prosecution.

Finally, one may argue that an anchoring effect is perfectly appropriate as many content elements correlate with unprotected speech and its ill

193. William N. Eskridge, Jr. & John Ferejohn, *Structuring Lawmaking to Reduce Cognitive Bias: A Critical View*, 87 CORNELL L. REV. 616, 635 (2002).

194. *See supra* notes 155-60 and accompanying text.

195. *See supra* notes 139-40 and accompanying text.

196. *See supra* notes 155-60 and accompanying text.

effects. One may argue, for example, that cross-burning tends to correlate with threatmaking in light of cross-burning's history and salience.[197] This point merits two responses. First, the fact that a content factor often correlates with an unprotected speech category is not a reason to encourage use of the former as a proxy for the latter. Indeed, free speech theory and doctrine suggest that one must remain highly wary of guesswork and imprecision in the realm of free speech law. Second, while wariness toward content-based regulations of unprotected speech is appropriate, such regulations are not categorically unconstitutional. Part V discusses the doctrinal standards that should be developed to balance wariness toward content-based regulations of unprotected speech against interests in permitting such regulations.

V. DOCTRINAL IMPLICATIONS: A CRITIQUE OF CURRENT DOCTRINE AND A PRESCRIPTION FOR NEW DOCTRINE

Part V considers the doctrinal implications of the theoretical analysis already undertaken. Subpart A explains the doctrinal standards that should apply to content-based regulations of unprotected speech. Subpart B contrasts Subpart A's approach with the doctrinal standards drawn by the *R.A.V.* Court, explaining that the *R.A.V.* standards are imprecise fits for containment-based concerns.

A. *The Right Doctrinal Standards*

Doctrinal standards to assess content-based regulations of unprotected speech must account for two factors. First, they must account for the virtues of categorization doctrine by according legislatures a fair amount of leeway to regulate unprotected speech categories. Second, to the extent that the standards tie legislatures' hands, they must do so in a manner responsive to the dangers that such regulations pose to free speech value and to protected speech.

From this perspective, two doctrinal standards seem appropriate. The first standard assesses the relationship between the content-based regulation and the harms targeted by the larger, unprotected speech category at issue. This standard should ask whether a content-based regulation of unprotected speech is substantially related to a state interest that constitutes a particularly compelling subset of the very reason the entire class of unprotected speech is proscribable. As Subpart B explains, this standard is similar to one of the several doctrinal standards invoked by the *R.A.V.* Court; indeed, it is similar to the only one of the Court's standards that approaches a reasoned response to containment-based

197. *See, e.g.*, Virginia v. Black, 538 U.S. 343, 353-57 (2003).

concerns. Yet as also discussed in Subpart B, the Court's version of this standard suffers from flaws that stem from the shakiness of the underlying theoretical foundations. Additionally, the Court's version of the standard is accompanied by a host of other standards that range from the superfluous to the affirmatively damaging. These differences are not only of theoretical import, but also have important practical consequences as illuminated in Part VI.

For now, it suffices to note that the standard suggested above strikes a balance between protecting free speech value and giving legislatures leeway to regulate subcategories of unprotected speech where justified. The standard protects free speech value in two ways. First, it accounts for the existence of free speech value within the realm of unprotected speech and the consequent dangers posed by content-based regulations by demanding that legislatures have compelling, nonsuperfluous reasons for content-based regulations. The compelling nature of the interest demanded is built into the standard. And superfluity is guarded against by demanding that the interest at issue be a particularly compelling subset, not a mere mirror image, of the reason why the larger speech category is regulated in the first place. Legislatures thus could not enact content-based regulations that target the precise harms already targeted by larger, unprotected speech categories. Furthermore, by demanding a "substantial relationship" to the harms, the standard guards against mere assertion or post hoc justification by demanding evidence that the content-based subcategory at issue raises real and uniquely pressing harms that were considered by the legislature. Second, the standard protects free speech value by minimizing the risk of protected speech infringement. Specifically, the standard serves as a "back-up check" of sorts against the risk that protected speech will be targeted or impacted[198] by assessing whether the harms targeted by the regulation are of the same nature as the harms targeted by the larger unprotected speech category. Finally, the standard strikes a balance between protecting free speech value and according legislatures leeway to regulate within categories of unprotected speech. Indeed, legislatures maintain substantial regulatory leeway that they do not possess in the realm of protected speech. To the extent that the harms that give rise to larger unprotected speech categories correlate with particular types of content, legislatures have leeway to create content-based regulations.

In addition to this standard, courts should consider more broadly whether protected speech is threatened in a manner that cannot be accounted for through a detailed standard. For example, in *Black*, as explained below, the Virginia statute not only prohibited threats conveyed

198. Such check is a "back-up check" in the sense that it ought to accompany, but of course not replace, any checks as to whether protected speech facially is infringed through a vague or overbroad regulation.

through cross-burning, but also made the fact of cross-burning prima facie evidence of a threat.[199] A court should recognize that the prima facie evidence provision suggests a desire to use the fact of cross-burning as a heuristic tool to indicate the presence of a threat and should invalidate the entire legislation as suspect in such a case.[200]

B. R.A.V. *and* Black*: Description and Critique of Existing Doctrine*

1. *R.A.V.*

In *R.A.V.*, the Court outlined a confusing series of standards that reflect the shakiness of their theoretical foundations.[201] First, the Court explained that content-based regulations of unprotected speech, like content-based regulations of protected speech, can survive if they pass strict scrutiny.[202] Yet without acknowledging that it is doing so, the Court effectively applies a much harsher version of strict scrutiny to the St. Paul ordinance than that which it typically applies to content-based regulations of protected speech.[203] In assessing content-based regulations of protected speech, the Court typically asks whether the regulation is narrowly tailored or is the least restrictive means of achieving a compelling government interest.[204] While the Court sometimes frames the inquiry by asking whether the regulation is a necessary means of achieving a compelling government interest,[205] the necessity inquiry typically focuses on the narrowness of the regulation. That is, the necessity inquiry typically focuses on whether the regulation is the least restrictive means of achieving a compelling government interest.[206] In contrast, the *R.A.V.* Court interprets the necessity inquiry to mean that a regulation will not be upheld if a broader, noncontent-based alternative will suffice.[207] The St. Paul ordinance fails to pass muster under this standard, because, the Court explains, a content-

199. *See infra* notes 242-43 and accompanying text.

200. *See infra* notes 262-65 and accompanying text.

201. R.A.V. v. City of St. Paul, 505 U.S. 377, 382-90 (1992).

202. *Id.* at 395-96.

203. *See id.* at 403-05 (White, J., concurring).

204. *See, e.g.*, United States v. Playboy Entm't Group, Inc., 529 U.S. 803, 813 (2000) (citing Reno v. ACLU, 521 U.S. 844, 874 (1997); Sable Communications of Cal., Inc. v. FCC, 492 U.S. 115, 126 (1989)).

205. *See, e.g.*, *R.A.V.*, 505 U.S. at 395 (citing Burson v. Freeman, 504 U.S. 191, 199 (1992) (plurality opinion); Perry Educ. Ass'n v. Perry Local Educators' Ass'n, 460 U.S. 37, 45 (1983)).

206. *E.g.*, *Burson*, 504 U.S. at 199-200, 206-08 (plurality opinion); *see also* Eberle, *supra* note 99, at 1162.

207. *R.A.V.*, 505 U.S. at 403-06 (White, J., concurring); Eberle, *supra* note 99, at 1144-45, 1161-63.

neutral prohibition of all fighting words would have sufficed to prohibit the particular fighting words singled out by the ordinance.[208]

The problem with the Court's use of what Edward J. Eberle calls heightened strict scrutiny[209] is that it seems effectively insurmountable and thus does no doctrinal work, making it at best superfluous and at worst a misleading rhetorical tool. The standard seems insurmountable because a content-neutral regulation of an entire category of unprotected speech by definition would eliminate the speech that a narrower regulation would eliminate. This is why the *R.A.V.* Court concludes that the St. Paul ordinance fails strict scrutiny, and the same would seem to be true of any content-based regulation of unprotected speech. If the standard does no doctrinal work, then at best it is a superfluous distraction. At worst, the standard is a misleading rhetorical tool, falsely bolstering the Court's "case" against a regulation that fails the Court's other, problem-fraught standards.

If the heightened strict scrutiny standard does no real doctrinal work, then any real doctrinal lifting is left to the *R.A.V.* Court's "exceptions" to its general rule against content-based regulations of unprotected speech.[210] Specifically, the *R.A.V.* Court explains that some content-based regulations simply pose no risk that "official suppression of ideas is afoot."[211] It outlines two specific exceptions in which no such risk exists, but also leaves a catch-all exception for any situation in which courts find that no such risk exists.[212] The catch-all exception is impossibly vague and underprotective of speech, leaving the door open for the Court to allow any regulation that strikes it as acceptable. The *R.A.V.* Court found that the St. Paul ordinance did not pass muster under this catch-all exception, partly because it did not pass muster under the Court's more specific exceptions and partly because the Court detected a viewpoint basis in the ordinance.[213]

As for the Court's more specific exceptions, one is for regulations that facially target speech content but that do so for purposes relating to the "secondary effects" of such speech, rather than to communicative or expressive aspects of the speech.[214] Secondary effects doctrine, itself a

208. *R.A.V.*, 505 U.S. at 395-96.

209. Eberle, *supra* note 99, at 1163, 1166.

210. *See R.A.V.*, 505 U.S. at 382-84.

211. *Id.* at 390.

212. *Id.* at 387-90.

213. *Id.* at 395.

214. *Id.* at 389; *see also* Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows*, *supra* note 1, at 1397 (discussing secondary effects doctrine). The "secondary effects" exception also is applied by the Court in the context of content-based regulations of protected speech, and the exception has been the subject of much judicial and scholarly commentary in that context. *See, e.g.*, Id. at 1397-98, 1398 n.280

deeply controversial area of free speech law, constitutes an exception to the presumption against content-based regulations where facially content-based regulations are underscored by a legislative purpose to prevent secondary effects, such as crime or violence, correlated with particular types of speech.[215] Although the *R.A.V.* Court deems this exception inapplicable to the St. Paul ordinance,[216] the Court's use of this exception is troubling for the same reason that it is troubling when it applies to protected speech. In the protected speech context, secondary effects doctrine is deeply controversial because a purported legislative wish to target the secondary effects of speech constitutes far too easy an excuse for otherwise impermissibly content-based regulation and because secondary effects are notoriously difficult to distinguish from primary emotive or communicative effects of speech.[217] In the context of unprotected speech, the fact that legislatures might purportedly wish to target secondary effects does not suffice to cure containment-based concerns regarding seemingly superfluous or otherwise suspicious legislation.[218]

---

215. *See* Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows*, *supra* note 1, at 1341 n.10, 1397-98; *see also infra* note 218 (discussing the *R.A.V.* Court's odd interpretation of secondary effects doctrine).

216. *R.A.V.*, 505 U.S. at 394. The Court explains that St. Paul deems the impact of fighting words on particularly vulnerable persons to constitute a secondary effect and that such impact in fact constitutes an emotive, or primary impact of fighting words. *Id.* The Court further explains that, had St. Paul argued that the particular fighting words at issue are more likely than other fighting words to provoke violent confrontations, such an argument, too, would rely on the words' emotive, primary impact on listeners, rather than on violence that simply happens to correlate with particular fighting words. *Id.* at 394 n.7.

217. Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows*, *supra* note 1, at 1397-98, 1398 n.280.

218. It also is worth noting that the Court's description of secondary effects doctrine is inconsistent with its articulation of that doctrine elsewhere and that this inconsistency likely is grounded in a desire to avoid the farthest reaching implications of *R.A.V.*'s general rule. As already noted, secondary effects doctrine typically allows facially content-based regulations to be treated as content-neutral when they are intended to target ill effects that happen to correlate with particular types of speech. *See supra* note 215 and accompanying text. As an example of secondary effects doctrine's application, however, the *R.A.V.* Court refers to words that are "swept up incidentally within the reach of a statute directed at conduct rather than speech," such as "sexually derogatory 'fighting words'" that "may produce a violation of Title VII's general prohibition against sexual discrimination in employment practices." *Id.* at 389. As Justice White (concurring with three other Justices in *R.A.V*) suggests, a content-based regulation of speech grounded in a purpose to prevent secondary effects is very different from a noncontent-based regulation targeting conduct that happens incidentally to include protected or unprotected speech. *Id.* at 409-11, 409 n.11 (White, J., concurring in the judgment). While the *R.A.V.* Court may well be correct that the latter is nonsuspect for any number of reasons, secondary effects doctrine is not one of those reasons. This aspect of the Court's opinion is noteworthy because, as the *R.A.V.* concurrence points out, it reflects some doctrinal flailing by the *R.A.V.* Court in its apparent effort to avoid the farthest reaching implications of its expansive general rule. *Id.* at 409 (White, J., concurring in the judgment).

The Court's more important exception—more important because it is a potentially large loophole and because it is similar to this Article's proposed standard—is its exception for situations in which "the basis for the content discrimination consists entirely of the very reason the entire class of speech at issue is proscribable."[219] The Court explains that:

> [s]uch 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. To illustrate: . . . . [T]he Federal Government can criminalize only those threats of violence that are directed against the President, since the reasons why threats of violence are outside the First Amendment (protecting individuals from the fear of violence, from the disruption that fear engenders, and from the possibility that the threatened violence will occur) have special force when applied to the person of the President. . . . But the Federal Government may not criminalize only those threats against the President that mention his policy on aid to inner cities.[220]

The Court explains that the St. Paul ordinance does not fall within this exception.[221] The Court says that fighting words are unprotected because they constitute a "particularly intolerable (and socially unnecessary) *mode* of expressi[on]," and that St. Paul has not singled out only fighting words that involve "an especially offensive mode of expression," but "has proscribed fighting words of whatever manner that communicate messages of racial, gender, or religious intolerance."[222]

To the extent that this standard attempts to link the means and ends of content-based regulations of unprotected speech to those of general regulations of unprotected speech, it might suggest a subtextual, containment-based concern on the Court's part. At the very least, it might suggest a concern with ensuring that only unprotected speech is targeted or impacted by such regulations. Not surprisingly, however, the murkiness of any containment-based concerns on the Court's part gives rise to serious problems in the Court's articulation and use of this standard.

There are two major problems with the Court's version of this exception, both as articulated and as applied. As articulated by the Court, the exception leaves open the possibility of regulations that are too

---

219. *Id.* at 388; *see also id.* at 408 (White, J., concurring in the judgment) (noting that this "exception swallows the majority's rule").

220. *Id.* at 388 (citation omitted).

221. *Id.* at 391.

222. *Id.* at 393-94.

superfluous to justify their risk to free speech value and to protected speech. The exception, on its face, allows for content-based subcategories of unprotected speech that respond to the precise harm already embraced by larger unprotected speech categories. Furthermore, even assuming that the exception is applied only to regulations targeting harms of the same nature as, but more compelling than, those targeted by the larger unprotected speech category,[223] the exception remains problematic because it demands no particular degree of fit between the regulation and the purported harms.

Perhaps most problematic is that the Supreme Court's interpretation of the exception in *R.A.V.* (and again in *Black*, as we shall see)[224] betrays uncertainty as to whether "the very reason [for proscribability]" refers to the harms targeted by the unprotected speech category or to some other characteristic shared by all speech within the category.[225] Such confusion, again, is not surprising in light of the tenuous justifications in which the exception is grounded in the first place. In *R.A.V.*, the Court adopts an interpretation that seems to hinge more on the nature of fighting words than on the harm that fighting words purport to target, explaining that fighting words constitute a "particularly intolerable (and socially unnecessary) *mode* of expressi[on]" and that the St. Paul ordinance does not single out fighting words expressed in a particularly offensive mode, but singles out fighting words reflecting particular topics or viewpoints.[226] Lost in this discussion is any serious consideration of whether the St. Paul ordinance adopts its content-distinction because the fighting words embraced by the distinction are particularly likely to cause immediate breaches of the peace or otherwise to cause the harms associated with fighting words.[227] As we shall see in Part VI, this has troubling implications for the types of regulations that the Court might be willing to uphold, particularly in the fighting words context.

Finally, the *R.A.V.* Court suggests that the St. Paul ordinance might be viewpoint-based, or at least might have a viewpoint-specific impact, and that this adds to its inherently suspicious nature.[228] While the implications of this observation—an observation with which other members of the Court vehemently disagree[229]—are left vague, the observation seems to

223. *See infra* text accompanying notes 244-45, 259 (explaining that the Court's application of this exception in *Black* is largely consistent with the "more compelling interest" interpretation of it).

224. *See infra* Part V.B.2.

225. *R.A.V.*, 505 U.S. at 388.

226. *Id.* at 393-94.

227. *Id.*

228. *See id.* at 391-92.

229. *Id.* at 433-35 (Stevens, J., concurring in the judgment).

color the Court's application of its standards. Indeed, the Court's questionable application of its "reason for proscribability" exception seems grounded partly in the Court's distaste for the fact that the ordinance, in the Court's view, singles out particular viewpoints.[230] The problem with the Court's emphasis on viewpoint-basis is that it is not well tailored to containment-based concerns. Because unprotected speech is unique in that its harms necessarily stem from speech content, it is not surprising that elements of speech content not normally proscribable—including viewpoint—may sometimes heighten the harms of the unprotected speech. Viewpoint elements thus might, at least in rare cases, justify content-based regulation of unprotected speech categories. On the other hand, content-based regulations based on subject matter or communicative symbols can pose substantial threats to free speech value that will not always prove justified by the requisite harms. The *R.A.V.* Court's fixation with viewpoint-basis thus threatens both to overprotect and to underprotect with respect to future regulations. Furthermore, it often is not clear whether a regulation is viewpoint-based, subject matter based, or based on communicative symbols or communicative impact. This is evidenced by the division within the *R.A.V.* Court as to whether the St. Paul ordinance indeed was viewpoint-based.[231] Such uncertainty, combined with the malleability of the *R.A.V.* standards, poses a risk that the Court will draw decisive conclusions about viewpoint-basis grounded in little more than a gut feeling about a regulation.

2. *Black*

This Subpart briefly sketches the facts of *Virginia v. Black*.[232] It then describes relevant parts of the *Black* Court's doctrinal analysis and critiques this analysis.

Two separate incidents gave rise to *Virginia v. Black*.[233] The first involved a Ku Klux Klan rally in which a twenty-five to thirty foot cross was burned.[234] The rally and cross-burning were held on private property adjacent to a state highway with the property owner's permission.[235] The rally and cross-burning were witnessed by a number of passerbys from the highway, including a sheriff.[236] After witnessing the cross-burning, the sheriff informed the rally's leader, Barry Black, that "'there's a law in the

230. *See infra* text accompanying note 232.

231. *See supra* notes 228-29 and accompanying text; *cf.* sources cited *supra* note 10 (discussing the difficulties in defining "content").

232. 538 U.S. 343 (2003).

233. *Id.* at 348-50.

234. *Id.* at 349.

235. *Id.* at 348.

236. *Id.*

State of Virginia that you cannot burn a cross'" and arrested him.[237] The second incident occurred when three men, including Richard Elliot and Jonathan O'Mara, burned a cross on the lawn of James Jubilee, an African-American man who recently had moved next door to Elliot.[238] The three men claimed to have burned the cross in order to "'get back' at Jubilee for complaining" to Elliot's mother about shots that Elliot fired in Elliot's backyard.[239]

Black, Elliot, and O'Mara each were charged with and convicted of violating a Virginia statute that provides:

> It shall be unlawful for any person or persons, with the intent of intimidating any person or group of persons, to burn, or cause to be burned, a cross on the property of another, a highway or other public place. . . .
>
> Any such burning of a cross shall be prima facie evidence of an intent to intimidate a person or group of persons.[240]

Black, Elliot and O'Mara each challenged the statute as facially unconstitutional.[241] The Virginia Supreme Court held in their favor, deeming the statute "'analytically indistinguishable from the ordinance found unconstitutional in *R.A.V.*'" because it punishes a content-based subcategory of threats.[242] The Virginia Supreme Court also held that the "prima facie evidence provision renders the statute overbroad because '[t]he enhanced probability of prosecution under the statute chills the expression of protected speech.'"[243]

The Supreme Court agreed that the statute is a content-based regulation of unprotected threats and that the principles of *R.A.V.* apply, but held that the statute falls within *R.A.V.*'s exception for situations in which "'the basis for the content discrimination consists entirely of the very reason the entire class of speech at issue is proscribable.'"[244] The *Black* Court concluded that Virginia may

> outlaw cross burnings done with the intent to intimidate because burning a cross is a particularly virulent form of intimidation. Instead of prohibiting all intimidating messages, Virginia may choose to regulate this subset of intimidating

237. *Id.* at 349.
238. *Id.* at 350.
239. *Id.*
240. *Id.* at 348 (quoting VA. CODE ANN. § 18.2–423 (Michie 1996)).
241. *Id.* at 351.
242. *Id.* (quoting Black v. Commonwealth, 553 S.E.2d 738, 742 (Va. 2001)).
243. *Id.* (quoting *Black*, 553 S.E.2d at 746).
244. *Id.* at 361 (quoting *R.A.V. v. City of St. Paul*, 505 U.S. 377, 388 (1992)).

> messages in light of cross burning's long and pernicious history as a signal of impending violence.[245]

The Court also emphasized that the regulation is not viewpoint- or subject-matter-based as it makes no difference under the statute whether speech is directed toward a "disfavored topic[]" and as "it is not true that cross burners direct their intimidating conduct solely to racial or religious minorities."[246]

The Supreme Court did, however, deem the statute's prima facie evidence provision unconstitutional.[247] The Court construed the prima facie evidence provision to mean that "'[t]he burning of a cross, by itself, is sufficient evidence from which [factfinders] may infer the required intent [to threaten].'"[248] The Court explained that the provision, so construed, creates a risk that cross-burning not intended to threaten and thus constitutionally protected will be punished.[249] Such punishment can occur where a defendant does not put on a defense and thus offers nothing to overcome the inference of intent to threaten.[250] Such punishment also can occur where, in close cases, factfinders treat the statutory inference as decisive.[251] The provision thus makes the statute overbroad, in the Court's apparent view.[252]

The *Black* Court's opinion includes a lengthy discussion of the history of cross-burning in the United States and of cross-burning's association with the Ku Klux Klan in particular.[253] The Court explains that a cross-burning can have one or both of two meanings. First, a cross-burning can be deeply expressive, albeit in a volatile and hurtful way, about prejudice or other divisive issues.[254] Second, a cross-burning can have the intent and

---

245. *Id.* at 363.

246. *Id.* at 362. The *Black* Court seems, in short, to deem the Virginia statute content-based by relying on a definition of "content" meaning the communicative manner or mode of expression, rather than viewpoint or subject matter. *See, e.g.*, Kitrosser, *From Marshall McLuhan to Anthropomorphic Cows*, *supra* note 1, at 1341-42 (defining communicative manner).

247. *Black*, 538 U.S. at 367.

248. *Id.* at 364. The Court based its interpretation on a Virginia state jury instruction. *Id.* (quoting jury instruction provided in Barry Black's case, which *Black* Court explains is the same as the Virginia Model Jury Instructions, Criminal, Instruction No. 10.250 (1998 and Supp. 2001)).

249. *Id.* at 366-67.

250. *Id.* at 365.

251. *Id.* at 366 (implying—though not stating explicitly—that the statute is overbroad); *see also id.* at 373 (Scalia, J., concurring in part, concurring in the judgment in part, and dissenting in part) ("In deeming § 18.2-423 facially invalid, the plurality presumably means to rely on some species of overbreadth doctrine.").

252. *Id.* at 364-67.

253. *Id.* at 352-57.

254. *Id.* at 357.

effect of placing persons in fear for their lives or physical safety.[255] The history and strength of the second meaning leads the Court to conclude that a cross-burning intended to intimidate constitutes a "particularly virulent form of intimidation," justifying an exception to *R.A.V.*'s general rule.[256] The existence of the first meaning leads the Court to conclude that each cross-burning must be examined separately for evidence of intent to threaten, lest a presumption of intent skew factfinder deliberations toward punishment in cases involving speech that not only is protected, but that constitutes "core political speech."[257]

The *Black* opinion reflects three major problems with existing doctrine. First, while the *Black* Court applies the "very reason [for proscribability]" exception[258] in a manner more appropriate than that of the *R.A.V.* Court, the contrast between the standard's application in the two cases highlights its troubling malleability. While the *Black* Court treats the standard as requiring statutes to target harms of the same nature but more compelling than those targeted by the larger unprotected speech category,[259] the standard continues, on its face, to leave room for statutes to target harms of the precise nature as those targeted by the larger unprotected speech category. More disturbing, the *Black* Court does nothing to reconcile its use of the standard with the *R.A.V.* Court's use of the same, leaving room for the Court to continue, in future cases, to use the standard to assess the nature of speech characteristics, rather than the harms targeted. This has troubling implications, as exemplified by the hypothetical fact pattern discussed in Part VI.D.

Second, the *Black* Court's emphasis on its perception that the statute is not viewpoint-based raises problems similar to those raised by the *R.A.V.* Court's emphasis on its perception that the St. Paul ordinance is viewpoint-based.[260] Assuming that it is easy to discern when a statute is based on viewpoint rather than subject matter or communicative symbol, the implication that the Court might apply its (already malleable) standards with less rigor in the latter cases is troubling. This is particularly so when one considers that the lines between the various types of content-distinctions are not so easy to discern, as reflected by a dispute between the Justices in *Black* as to whether the statute truly is viewpoint-neutral.[261]

255. *Id.*
256. *Id.* at 363.
257. *Id.* at 365.
258. R.A.V. v. City of St. Paul, 505 U.S. 377, 388 (1992).
259. *Supra* notes 244-45 and accompanying text.
260. *See supra* notes 228-31 and accompanying text.
261. *Compare Black*, 538 U.S. at 361-62, *with id.* at 383-84 (Souter, J., concurring in the judgment in part and dissenting in part).

Third, the *Black* Court's treatment of the prima facie evidence provision, while a step in the right direction, suggests that the Court does not fully grasp the nature of the heuristic skewing risk posed by content-based regulations of unprotected speech. The question does not hinge, as the majority implies and partial dissenters assume, on whether the statute is facially overbroad as a result of the prima facie evidence provision.[262] Of course, facial overbreadth is an important factor to consider. However, even where a statute is not overbroad in that it only targets unprotected speech on its face, a risk of heuristic skewing can remain and a prima facie evidence provision can remain suspicious from this perspective. Furthermore, as noted above[263] and explained further below,[264] the nature of a heuristic skewing risk is such that the prima facie evidence provision should have led the Court to invalidate the entire statute, not just the prima facie evidence provision by itself.[265]

VI. SOME THOUGHTS ON APPLICATION: APPLYING THIS ARTICLE'S STANDARDS TO THE FACTS OF *R.A.V.* AND *BLACK* AND TO TWO HYPOTHETICAL STATUTES

Subparts A and B apply this Article's standards to the provisions at issue in *R.A.V.* and *Black*. While the result reached in *R.A.V.* would likely have been the same under this Article's approach and the result in *Black* would have been partly the same, there would have been significant differences in rationale with important consequences for future cases. These consequences are reflected in the comparative application of this Article's standards and the *R.A.V.* Court's standards to two hypothetical statutes discussed in Subparts C and D.

---

262. *See Black*, 538 U.S. at 366; *id.* at 371-73 (Scalia, J., concurring in part, concurring in the judgment in part, and dissenting in part).

263. *See supra* note 200 and accompanying text.

264. *See infra* note 286 and accompanying text.

265. *See Black*, 538 U.S. at 386 (Souter, J., concurring in the judgment in part and dissenting in part). Justice Souter stated,

> [t]he question here is not the permissible scope of an arguably overbroad statute, but the claim of a clearly content-based statute to an exception from the general prohibition of content-based proscriptions, an exception that is not warranted if the statute's terms show that suppression of ideas may be afoot. Accordingly, the way to look at the prima facie evidence provision is to consider it for any indication of what is afoot.

Id. (Souter, J., concurring in the judgment in part and dissenting in part).

A. R.A.V. *Revisited*

Assuming arguendo that the St. Paul ordinance at issue in *R.A.V.* was not facially overbroad,[266] the regulation should have been struck down in light of the failure of St. Paul's justifications to demonstrate that the ordinance was "substantially related to a government interest that constitutes a particularly compelling subset of the very reason the entire class of unprotected speech is proscribable."[267] Judging by St. Paul's brief to the U.S. Supreme Court and by the Minnesota Supreme Court's opinion, the ordinance appears to have been justified by a loose mix of factors, but primarily by the notion that the speech targeted conveys a message of group hatred and that "[i]t is the responsibility, even the obligation, of diverse communities to confront such notions in whatever form they appear."[268] The most obvious problem with this goal is that it does not relate directly to the purposes underlying fighting words provisions generally. While there is some dispute as to the definition of fighting words,[269] a fairly clear meaning of the term is that it refers at least to those words "reasonably considered direct personal insults likely to invite physical confrontation,"[270] and that it thus is linked to an interest in stopping imminent breaches of the peace.[271] A desire to send an anti-group-hatred message simply does not constitute a "more compelling subset" of this interest.

While St. Paul's brief refers at points to the particularly deep-seated and sustained fears to which cross-burning gives rise in members of racial

---

266. *See supra* notes 58-59 and accompanying text; *infra* note 277 and accompanying text.

267. *See supra* Part V.A (proposing this standard).

268. *In re* Welfare of R.A.V., 464 N.W.2d 507, 508 (Minn. 1991); *see also* Brief for Respondent at 7, R.A.V. v. City of St. Paul, 505 U.S. 377 (1992) (No. 90-7675). Respondent's Brief further stated

> there is a need for specific enactments to deal with the use of symbols of racial, religious, ethnic or other group hatred. Otherwise, there is a danger that minority groups will not understand that this aspect of the behavior they have experienced is not condoned by the majority. This is a lesson the majority also must understand and openly acknowledge.

Id. at 25.

269. This dispute is hardly advanced by the *R.A.V.* Court's insistence on relying, at least arguendo, on the Minnesota Supreme Court's "limiting construction" of the St. Paul ordinance, which itself seems to exceed the modern definition of fighting words by stating that "the ordinance censors only those displays that one knows or should know will create anger, alarm or resentment based on racial, ethnic, gender or religious bias." *Welfare of R.A.V.*, 464 N.W.2d at 510; *see also R.A.V.*, 505 U.S. at 381 (deferring to the state court's construction). *But see* Shiffrin, *supra* note 120, at 72-73 (defending state court's construction).

270. *See supra* notes 33, 135 and accompanying text.

271. *See supra* notes 33, 135 and accompanying text.

minority groups,[272] this interest, too, fails to justify St. Paul's ordinance upon close analysis. First, the interest is presented largely as an afterthought to the goal of making a statement against group hatred and even is conflated at points with an interest in suppressing protected speech. Indeed, St. Paul's brief at one point discusses the anti-fear interest interchangeably with an interest in countering all racial insults and epithets.[273] More convincing evidence of an anti-fear interest unconnected to a desire to suppress protected speech is necessary to prove a substantial relationship between such interest and the regulation. Second, St. Paul's references to cross-burning and race-based fears have little bearing on the facial validity of the statute as a whole given the statute's emphasis on gender and religion, as well as race, and given the statute's application to swastikas and other "symbol[s], object[s], appellation[s], characterization[s and] graffiti."[274] If a "substantial relationship" between means and ends is demanded to guard against purposeful or inadvertent infringement on protected speech, then a justification underlying one prong of a multi-pronged statute cannot suffice to justify the statute as a whole. Finally, while an interest in stemming particularly sustained and deep-seated fears may constitute a more compelling subset of an interest directed toward stemming fear generally, fear-reduction is not properly conceptualized as an interest underscoring fighting words doctrine. While there is controversy over the scope of fighting words doctrine, as already noted, and while early case law provides a basis for argument to the effect that fighting words include words that cause fear and other deep-seated emotional harms,[275] recent case law strongly suggests that fighting words encompass only those words that threaten an imminent breach of the peace.[276] Assuming, then, that fighting words encompass only speech causing imminent peace breaches, an interest in stemming fear-based harms simply does not constitute a "more compelling subset" of the interest underscoring fighting words doctrine.

These concerns are compounded by the ordinance's questionable breadth. While the *R.A.V.* Court avoided deciding whether the ordinance was facially overbroad, it is relevant, from a containment-based perspective, that the content-based ordinance accompanied a vague and

---

272. *See, e.g.*, Brief for Respondent at 26 n.5, 36 n.5, 38-41, *R.A.V.* (No. 90-7675) (citing testimony of the fears cross-burning generates in the African-American community).

273. *Id.* at 37-40; *cf. id.* at 19 (suggesting that cross-burning generally has little or no speech value and thus can readily be punished). The anti-fear interest's status as an apparent afterthought also is reflected in the fact that some of the Brief's most compelling references to the interest are conveyed entirely through footnotes. *See id.* at 26 n.5, 36 n.5.

274. *R.A.V.*, 505 U.S. at 380.

275. *See* Gard, *supra* note 33, at 531-34 (citing Chaplinsky v. New Hampshire, 315 U.S. 568, 571-72 (1942)).

276. *See supra* note 134 and accompanying text.

facially broad description of fighting words. Such accompaniment bolsters the concern that St. Paul sought to use fighting words doctrine as a legal hook to punish protected racist speech.[277]

While this Article's approach leads to the same result reached by the *R.A.V.* Court, the underlying rationale to which the approach lends itself stands on much firmer theoretical footing and offers sounder doctrinal guidance than does the rationale offered by the *R.A.V.* Court. This Article's approach involves application of a more rigid and more carefully justified version of the "reason for proscribability" standard, as detailed above.[278] Additionally, peripheral inquiries that may unduly restrain legislatures where their provisions are deemed viewpoint-based or that may underprotect speech by too readily permitting provisions based on communicative mode are avoided. Furthermore, this Article's approach, which joins the "more compelling subset" question to a "substantial relationship" requirement but does not pretend that standard strict scrutiny is relevant to the inquiry, avoids the difficulties faced by the *R.A.V.* Court as it struggled to explain why the St. Paul ordinance could not meet strict scrutiny.[279]

B. Black *Revisited*

Thus, while St. Paul's ordinance fails to pass muster under this Article's doctrinal approach, both the approach and St. Paul's discussion of the unique, race-based fears caused by cross-burning lend themselves to a principled explanation for the result in *Virginia v. Black* to the extent that the Virginia statute's general prohibition is considered apart from its prima facie evidence provision. As St. Paul's analysis, Virginia's analysis in its briefs to the Supreme Court in *Black,* and the Supreme Court's discussion in *Black* suggest, there is much evidence to the effect that cross-burning creates paralyzing fear and intimidation of a uniquely sustained and deep-seated nature.[280] The evidence suggests both that cross-burning may more readily create fears than do other threats and also that the fears caused themselves are uniquely profound.[281] Thus, a statute directed toward threatening cross-burnings may well pass muster as being

277. *See R.A.V.*, 505 U.S. at 411-15 (White, J., concurring in the judgment) (deeming the St. Paul ordinance overbroad).

278. *See supra* Part V.A.

279. *See supra* Part V.B.1.

280. *See, e.g.*, Virginia v. Black, 538 U.S. 343, 352-59 (2003) (describing the intimidation created by Klan cross-burnings); Brief of Petitioner at 29, 33-37, Virginia v. Black, 53 U.S. 343 (2003) (No. 01-1107) (describing society's understanding of cross-burning); Brief for Respondent at 26 n.5, 36 n.5, 38-41, *R.A.V.* (No. 90-7675) (highlighting the historical association of cross-burning with violence).

281. Brief for Respondent at 26 n.5, 36 n.5, 38-41, *R.A.V.* (No. 90-7675).

substantially related to a more compelling subset of the interests deemed to underscore threats provisions generally.[282]

Removing the prima facie evidence provision from the picture, then, and assuming that this Article's approach would lead to the same result as that reached in *Black*, the approach would involve an important difference in rationale. Specifically, the result would be grounded solely in the relationship between the state interests underscoring threat provisions generally and the interests underscoring the cross-burning threats provision at issue. So long as the latter constitute a more compelling subset of the former and relate substantially to the statute at issue, the provision can be deemed a legitimate unprotected speech provision rather than a threat to protected speech. Irrelevant to this inquiry are attempts to discern whether the state interests are based on viewpoint, subject matter, or communicative mode. The danger that such attempts will diminish protection for symbolic speech by minimizing the significance of infringements thereupon is exemplified by Virginia's brief, which repeatedly argues that the relevant provision is not content-based at all because it merely targets one communicative mode of expression.[283] And the confusion and disingenuousness to which such attempts lend themselves is exemplified by the *Black* Court's attempts to make clear that the fears caused by threatening cross-burnings are not unique to particular racial or religious groups,[284] presumably because such uniqueness might signal a statutory viewpoint or subject matter basis.[285]

Finally, under the containment-based scrutiny required under this Article's approach, much greater attention should be paid to the relevance of the Virginia statute's prima facie evidence provision to the soundness of the statute as a whole. As Justice Souter indicates in partial dissent, the prima facie evidence provision suggests that the statute, if not as initially created then as maintained, is intended at least partly to "skew jury

282. *See, e.g.*, Rothman, *supra* note 76, at 285 (discussing policies behind threats doctrine).

283. Brief of Petitioner at 11-21, *Black* (No. 01-1107).

284. *See Black*, 538 U.S. at 352-56, 362 (listing examples of cross-burning used to target various racial, religious, and labor groups).

285. A description of cross-burning as causing uniquely paralyzing harms for particularly vulnerable groups is fully consistent with the conclusion that threatening cross-burnings cause a more compelling subset of the harms caused by threats generally. Nonetheless, the *Black* Court seemed to make an effort to distance itself from this conclusion. *See id.* at 365-67. Professor Richard Delgado, commenting in the *New York Times* after the oral argument in *Black*, cited the strained content-neutrality arguments made by the statute's proponents as a basis for criticizing the content-distinction generally. Adam Liptak, *Justice Thomas's Stand; Symbols and Free Speech*, N.Y. TIMES, Dec. 15, 2002, § 4 (Week in Review), at 5 (quoting Professor Delgado as follows: "'Burning crosses are expressions of bigotry associated with racism in the South,' . . . and that message is precisely why they should be banned. The requirement of content neutrality is . . . 'a mechanistic test' that dishonors history and reality").

deliberations toward conviction in cases where the evidence of intent to intimidate is relatively weak and arguably consistent with a solely ideological reason for burning."[286] This conclusion, and consequent invalidation of the entire statute, is appropriate from a containment-based perspective.

C. *A Hypothetical Race-Based Threats Statute*

The practical significance of this Article's approach is exemplified by comparing how a hypothetical statute punishing race-based threats might fare under this Article's approach with the statute's likely fate under the *R.A.V.* Court's approach. It is very likely that such a statute would pass muster under this Article's approach. If Virginia could demonstrate that threats that evince race-based motivation have a history of inspiring particularly great fear and paralysis because of the unique likelihood that they will be carried out and that such concerns motivated the state legislature, such a statute likely would pass muster under a containment-based approach. Under the Supreme Court's approach, however, the result would be far less clear. On one hand, the Court might interpret its "reason for proscribability" exception[287] in a manner consistent with that just described. On the other hand, the Court might repeat three facets of its performance in *R.A.V.* to strike the statute down. First, it might rely on the statute's arguably viewpoint-based nature[288] to deem it unconstitutional. Second, the Court could rely on the malleability of its "reason for proscribability" exception and apply the exception to assess something other than the targeted harms. The Court could thus conclude that the statute does not fall within the exception. The Court likely would combine this discussion with an emphasis on the statute's viewpoint-basis, much as the Court did in *R.A.V.*[289] Third, the Court could bolster its analysis rhetorically by invoking other doctrinal elements, particularly the heavy hand of heightened strict scrutiny.

---

286. *Black*, 538 U.S. at 383-85 (Souter, J., concurring in the judgment in part and dissenting in part).

287. *See supra* notes 219-27 and accompanying text.

288. It is not at all clear that this hypothetical statute is viewpoint-based rather than subject-matter-based or communicative-impact-based. However, the debate between the *R.A.V.* Justices as to the nature of the St. Paul ordinance, *see supra* notes 228-31 and accompanying text, suggests that there would be judicial debate, at the very least, as to the nature of this hypothetical statute. *Compare* R.A.V. v. City of St. Paul, 505 U.S. 377, 391 (1992) ("In its practical operation, moreover, the ordinance goes even beyond mere content discrimination, to actual viewpoint discrimination."), *with id.* at 433 (Stevens, J., concurring in the judgment) ("Significantly, the St. Paul ordinance regulates speech not on the basis of its subject matter or the viewpoint expressed, but rather on the basis of the *harm* the speech causes.").

289. *See supra* notes 228-31 and accompanying text.

D. *A Hypothetical Fighting Words Statute*

Just as the Court could rely on its doctrine's malleability to strike down a race-based threats statute, so the Court could rely upon such malleability to uphold statutes that would be struck down under a containment-based approach. One suggestive example is a hypothetical statute punishing fighting words that involve profane language. Under a containment-based approach, such a statute likely would be struck down for at least one of two reasons. First, while it would not be surprising to learn that profanity is one of many factors sometimes encompassed in fighting words, it seems unlikely that profanity would have a provably heightened tendency to create or exacerbate physical fights above and beyond the tendencies of fighting words generally to have such effect. Second, from a perspective focused on the relationship between the harms of the larger unprotected speech category and the harms of the content-based subcategory thereof, it is possible that fighting words are categorically unlikely to be amenable to constitutional, content-based subcategorization. It is the very nature of fighting words that they are exceptional compared to merely provocative or upsetting language, rising effectively to the level of face-to-face invitations to fight. Thus, it would be uniquely difficult for subcategories of fighting words to rise beyond this already exceptional level of discourse to constitute words that are particularly likely to spark fights or to spark particularly bloody fights. In any event, whether based on a case-specific application of this Article's standard or a more categorical assessment of this Article's standard in relation to fighting words, it seems highly likely that this hypothetical statute would be struck down under a containment-based approach.

There is a good chance, however, that this hypothetical statute would be upheld under the *R.A.V.* doctrine. First, the last time that the Court applied its "reason for proscribability" exception in the context of fighting words—namely, in the *R.A.V.* case—the Court deemed the relevant question whether the subcategory at issue constitutes a "particularly intolerable (and socially unnecessary) *mode* of expressi[on]."[290] It is entirely plausible that the Court would deem profane fighting words a "particularly intolerable (and socially unnecessary) *mode* of expressi[on]."[291] Second, this hypothetical statute would give the Court

290. *See supra* note 226 and accompanying text.

291. The Supreme Court's treatment of profane language has not been entirely consistent over the years. *Compare, e.g.*, Cohen v. California, 403 U.S. 15, 15-17 (1971) (invalidating misdemeanor conviction for wearing a jacket bearing the words "Fuck the Draft" in a courthouse corridor), *with* FCC v. Pacifica Found., 438 U.S. 726, 729-33, 750 (1978) (upholding FCC fine for radio broadcast of George Carlin's "Filthy Words" monologue and explaining that the FCC may require radio stations to channel "indecent" broadcasts to times when children are unlikely to be in the listening audience). Nonetheless, the Court has evinced relative deference toward profanity

another opportunity to invoke its relative lack of suspicion of content-based regulations that are based on communicative symbol, manner, or mode rather than viewpoint. This point could be of independent significance or of rhetorical assistance in bolstering a lenient application of its standards. In either case, the Court's approaches in *R.A.V.* and *Black* suggest a likelihood that the Court would uphold this hypothetical statute.

VII. CONCLUSION

While the Supreme Court surely is "onto something" in its instinctive suspicion of content-based regulations of unprotected speech, precedent suffers badly for the Court's failure to grasp the nature of that "something." The venture is not, however, unsalvageable. Should the Court acknowledge containment-based concerns as the basis for suspicion over content-based regulations of unprotected speech, the Court could breathe new life into this area of the law.

Identifying the categorization paradox and the containment principle as the analytical cores of an approach to content-based regulations of unprotected speech is important on theoretical, doctrinal, and practical levels. As a theoretical matter, the precedent as it currently stands evinces inattentiveness at best and misunderstanding at worst of the value and implications of categorization doctrine and the relationship between that doctrine and content-distinction principles. As a doctrinal matter, such theoretical difficulties naturally translate into problems in formulating standards.

On a practical level, adjusting the theory and doctrine of the "*R.A.V.* problem" to conform to the containment principle is highly significant, even in cases where it is not outcome-determinative. In such cases, problematic theory and doctrine shape precedent for future cases, future litigation strategies, and future legislative drafting. Furthermore, to the extent that the case law seems confused or disingenuous, there is the cost of diminished public faith in the judiciary. Moreover, replacing the theory and doctrine of *R.A.V.* with a containment-based approach often will be outcome-determinative, as the hypothetical statutes in Part VI demonstrate.

In short, the doctrine and theory of content-based regulations of unprotected speech are as significant as they are confounding. Embracing the containment principle by no means solves all problems, as difficult

---

and indecency regulations that are fairly narrow in nature (e.g., that narrow their scope by requiring only time channeling). *Pacifica Found.*, 438 U.S. at 750. This point, combined with the *R.A.V.* Court's dicta suggesting a favorable view of restrictions on "particularly intolerable" fighting words, *see supra* notes 221-22, 226 and accompanying text, suggests that a "profane fighting words" regulation likely would be upheld.

questions of application remain. Such embrace is, however, an important step toward a coherent approach to the problem. Such embrace also has broader potential to facilitate understanding of the worth and the dangers of categorization doctrine and the need to keep such doctrine vibrant within its domain, but carefully contained.