

HATE SPEECH AND IDENTITY POLITICS: A SITUATIONALIST PROPOSAL

JULIE SEAMAN*

I. INTRODUCTION.....	99
II. HATE SPEECH AND THE UNIVERSITY.....	105
III. SOCIAL PSYCHOLOGY: INDIVIDUALS, GROUPS, AND BEHAVIOR.....	112
A. <i>Deindividuation Theory and the Psychology of Anonymity</i>	112
B. <i>Social Identity and Priming</i>	116
IV. CONCLUSION: THE POWER OF THE SITUATION AND THE POWER OF THE GROUP	121

I. INTRODUCTION

The scholarly debate over campus hate speech codes is most often characterized as a clash of absolutes, a conflict between two irreconcilable moral and political visions. On one side are the so-called “free speech absolutists,” who reject hate speech¹ restrictions on campuses and elsewhere based on their incompatibility with fundamental precepts of liberal democracy and individual autonomy.² On the other side are critical race theorists and antipornography feminists who ar-

* Associate Professor, Emory University School of Law. J.D., Harvard Law School. I wish to thank the participants in the 2008 First Amendment Discussion Forum for stimulating conversation and thoughtful comments on an earlier draft of this Essay. In particular, Russell Weaver’s energy and enthusiasm were instrumental in making the Forum a wonderfully constructive experience.

1. “‘Hate speech’ is the generic term that has come to embrace the use of speech attacks based on race, ethnicity, religion, and sexual orientation or preference.” Rodney A. Smolla, *Academic Freedom, Hate Speech, and the Idea of a University*, 53 LAW & CONTEMP. PROBS. 195, 195 (1990). However, the definition of hate speech is by no means settled or clear. See, e.g., Mari J. Matsuda, *Public Response to Racist Speech: Considering the Victim’s Story*, 87 MICH. L. REV. 2320, 2361-70 (1989) (considering several “hard cases” that fall at the edges of her definition of hate speech); Ronald Turner, *Regulating Hate Speech and the First Amendment: The Attractions of, and Objections to, an Explicit Harms-Based Analysis*, 29 IND. L. REV. 257, 257-58 n.1 (1995) (collecting several different definitions of hate speech). For the purposes of this Essay, “I know it when I see it” should be sufficient to define the type of speech to which these ideas are directed. Cf. *Jacobellis v. Ohio*, 378 U.S. 184, 197 (1964) (Stewart, J., concurring) (“I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description [of hard-core pornography that may be criminalized under the First and Fourteenth Amendments]; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it, and the motion picture involved in this case is not that.”).

2. As Professor Steven Gey has pointed out, the “absolutist” label is largely a rhetorical device rather than an accurate description of the position. See Steven G. Gey, *The Case Against Postmodern Censorship Theory*, 145 U. PA. L. REV. 193, 219-20 n.84 (1996); see also Nadine Strossen, *Regulating Racist Speech on Campus: A Modest Proposal?*, 1990 DUKE L.J. 484, 490 (disagreeing with the characterization of the civil libertarian position on hate speech, which was taken by a prominent critical race scholar, and specifically stating that “contrary to Professor [Charles] Lawrence’s assumption, traditional civil libertarians do not categorically reject every effort to regulate racist speech”). Recognizing this, I have put the descriptive label in quotation marks. I mention the extreme descriptions in order to frame the debate between what are in fact much more complicated and nuanced positions on both sides that, nonetheless, appear to be fundamentally irreconcilable.

gue that hate speech both creates and perpetuates a poisonous social atmosphere in which minorities and women are unable to realize genuine equality of citizenship.³ Too often, and ironically in the context of a debate about the appropriate contours of the First Amendment's protection of freedom of speech, it seems as though the two sides are talking past one another rather than engaging in anything resembling a constructive dialogue.

It may be that there is no way to reconcile these polar positions, no middle ground that can be true to the assumptions that drive either side of the hate speech debate.⁴ As proponents of hate speech restrictions argue, hate speech potentially causes great harm;⁵ as opponents point out, government restriction of hate speech is content (and arguably viewpoint) based and is a form of state-imposed political orthodoxy.⁶ Yet, despite an apparent consensus among the reported judicial decisions that sweeping hate speech regulations at public universities violate the First Amendment, issues regarding the legitimacy and desirability of campus hate speech restrictions are unlikely to disappear from the scholarly and political landscape.⁷

3. See, e.g., Richard Delgado, *Words That Wound: A Tort Action for Racial Insults, Epithets, and Name-Calling*, 17 HARV. C.R.-C.L. L. REV. 133, 181 (1982) [hereinafter Delgado, *Words That Wound*] (concluding that a "review of the social science literature revealed that racism and racial insults influence the parenting practices of minority individuals and have a very great effect on children, thus perpetuating the harms of racism"); Matsuda, *supra* note 1, at 2339 ("[A]t some level, no matter how much both victims and well-meaning dominant-group members resist it, racial inferiority is planted in our minds as an idea that may hold some truth."); Laura Beth Nielsen, *Subtle, Pervasive, Harmful: Racist and Sexist Remarks in Public as Hate Speech*, 58 J. SOC. ISSUES 265, 266 (2002) ("If prejudice is about relative group position, then public hate speech provides a clear example of one of the ways in which such social hierarchies are constructed and reinforced on a day-to-day basis.").

4. See Richard Delgado, *Campus Antiracism Rules: Constitutional Narratives in Collision*, 85 NW. U. L. REV. 343, 346 (1991) (describing intractability of hate speech controversy as a framing issue and arguing that "[t]he legal analysis, therefore, leads to opposite conclusions depending on the starting point[, wherein] . . . both sides invoke different narratives to rally support"); see also Kenji Yoshino, *The Eclectic Model of Censorship*, 88 CAL. L. REV. 1635, 1636-38 (2000) (reviewing CENSORSHIP AND SILENCING: PRACTICES OF CULTURAL REGULATION (Robert C. Post ed., 1998)) (contrasting the "classical model" of censorship with the "Foucaultian model" and then describing an intermediate position which he labels "eclectic").

5. See, e.g., RICHARD DELGADO & JEAN STEFANCIC, UNDERSTANDING WORDS THAT WOUND 11-18 (2004) (summarizing harms of hate speech, including physical, psychological, economic, and social harms); Delgado, *Words That Wound*, *supra* note 3 (proposing a tort action for racial insults and examining the harms caused by such insults); Charles R. Lawrence III, *If He Hollers, Let Him Go: Regulating Racist Speech on Campus*, 1990 DUKE L.J. 431 (arguing that protection of speech entails balancing of harms and offering a personal anecdote to illuminate the nature and extent of the harms of racist speech).

6. See *R.A.V. v. City of St. Paul*, 505 U.S. 377, 392 (1992) (classifying the city's hate speech ordinance as viewpoint discrimination and stating that "[t]he point of the First Amendment is that majority preferences must be expressed in some fashion other than silencing speech on the basis of its content").

7. Federal courts have invalidated every university hate speech code that has thus far been challenged, primarily on grounds of overbreadth and vagueness. See, e.g., *Coll. Republicans at San Francisco State Univ. v. Reed*, 523 F. Supp. 2d 1005, 1021, 1024 (N.D. Cal.

In this Essay, I consider whether recent scientific findings about the nature of human attitude formation and decisionmaking, and the social cues that drive behavior, might have the potential to shift the hate speech debate such that some areas of common ground come into view. Specifically, new insights in the areas of social and cognitive psychology and brain development suggest that there may be a “third way” that escapes some of the most difficult criticisms of both the strong free speech and the procensorship positions and that incorporates some of the most important insights of each.

Scholars who argue in favor of hate speech restrictions have been characterized as strong social constructionists⁸ who believe that group-based identities—for example, race and gender—arise not from any inherent traits of individuals but rather from social meanings that are created by the words and actions of others.⁹ Thus, critical race theorists generally assert that racist hate speech, along with other social and cultural mechanisms, works to construct minorities as inferior and subordinated, thus launching a self-fulfilling cycle of discrimination and inequality.¹⁰ Similarly, antipornography feminists in

2007) (granting preliminary injunction, on First Amendment overbreadth grounds, against enforcement of university speech code provisions that empowered the university to punish students for behavior that was not “civil” or was “inconsistent with [the university’s] goals”); *Roberts v. Haragan*, 346 F. Supp. 2d 853, 872-73 (N.D. Tex. 2004) (holding that university speech code was overbroad); *Bair v. Shippensburg Univ.*, 280 F. Supp. 2d 357, 372-73 (M.D. Pa. 2003) (granting preliminary injunction in favor of student challenging university speech code based on finding that code was facially overbroad); *Dambrot v. Cent. Mich. Univ.*, 39 F. Supp. 477, 482-84 (1993) (holding speech code, under which basketball coach was disciplined for using a racial epithet, violated the First Amendment); *UWM Post, Inc. v. Bd. of Regents of Univ. of Wis. Sys.*, 774 F. Supp. 1163, 1178-81 (E.D. Wis. 1991) (finding university rule prohibiting discriminatory epithets overbroad and unduly vague); *Doe v. Univ. of Mich.*, 721 F. Supp. 852, 866-67 (E.D. Mich. 1989) (holding university antidiscrimination code unconstitutional). Nonetheless, institutions continue to maintain their codes and, in many cases, to strengthen them. See Jon B. Gould, *The Precedent That Wasn’t: College Hate Speech Codes and the Two Faces of Legal Compliance*, 35 LAW & SOC’Y REV. 345, 345, 357-59 (2001) [hereinafter Gould, *College Hate Speech Codes*] (finding, in an empirical investigation, that many colleges and universities—both public and private—continue to enact or maintain hate speech codes despite court decisions striking down similar codes as unconstitutional); see also JON B. GOULD, SPEAK NO EVIL: THE TRIUMPH OF HATE SPEECH REGULATION (2005) (examining in detail the implications of his empirical findings). In addition, because private universities are not subject to First Amendment constraints, they have room to experiment in the delicate balancing between freedom of speech and promotion of equality.

8. See Gey, *supra* note 2, at 198 (arguing that “[t]he social constructionist argument is perhaps the clearest thread linking the various groups proposing new theories to justify speech regulation”).

9. See, e.g., IAN HACKING, THE SOCIAL CONSTRUCTION OF WHAT? 1 (1999) (containing an alphabetical listing of traits and other things said to be socially constructed, including gender, facts, illness, reality, and serial homicide). For an account of the social constructionist position in the context of gender, see Julie A. Seaman, *Form and (Dys)Function in Sexual Harassment Law: Biology, Culture, and the Spandrels of Title VII*, 37 ARIZ. ST. L.J. 321, 355-61 (2005).

10. See DELGADO & STEFANCIC, *supra* note 5, at 26-27 (“Words play a central role in the creation of all these hurdles [that minorities endure in society] [I]t is racial depic-

the tradition of Catharine MacKinnon argue that gender is socially constructed and that violent or degrading sexual depictions both create and perpetuate a society in which individuals understand women as objects of male sexual desire and agency.¹¹

In contrast, traditional free speech theorists tend to reject this vision of false consciousness on the part of individual citizens who cannot know their “true” desires or beliefs.¹² Indeed, all of the dominant theories of the First Amendment rest upon the assumption of an autonomous subject who is free to weigh and choose among competing ideas in some nontrivial manner.¹³

Notwithstanding this profound disagreement about the respective roles of the individual and of society in shaping identity, I suspect neither side would quarrel with the commonplace observation that indi-

tion and imagery that maintain the system of white-over-black in this society, now that those more tangible barriers [of slavery and Jim Crow] have been lowered.”); *see also* JUDITH BUTLER, *EXCITABLE SPEECH: A POLITICS OF THE PERFORMATIVE* 5 (1997) (“[I]t is by being interpellated within the terms of language that a certain social existence of the body first becomes possible.”).

11. *See* CATHARINE A. MACKINNON, *ONLY WORDS* 13 (1993) (“Social inequality is substantially created and enforced—that is, *done*—through words and images Elevation and denigration are all accomplished through meaningful symbols and communicative acts in which saying it is doing it.”). MacKinnon and other radical feminists have defined “pornography” very specifically as both broader and narrower than the unprotected category “obscenity.” Under the proposed antipornography ordinance drafted by MacKinnon and Andrea Dworkin, pornography is defined as:

[T]he graphic sexually explicit subordination of women, whether in pictures or in words, that also includes . . . [w]omen [who] are presented as[:] sexual objects who enjoy pain or humiliation; . . . experience sexual pleasure in being raped; . . . tied up or cut up or mutilated or bruised or physically hurt[:] . . . being penetrated by objects or animals; . . . and . . . in scenarios of degradation, injury, abasement, torture, shown as filthy or inferior, bleeding, bruised, or hurt in a context that makes these conditions sexual

Am. Booksellers Ass’n v. Hudnut, 771 F.2d 323, 324 (7th Cir. 1985) (quoting Indianapolis Code § 16-3(q)). MacKinnon views pornography, as defined in the model antipornography ordinance, as a variety of hate speech against women. *See, e.g.*, MACKINNON, *supra*, at 104 (“Hate speech and pornography do the same thing: enact the abuse.”). In contrast, the Supreme Court’s definition of obscenity, which is unprotected speech under First Amendment doctrine, turns not on its depiction of subordination but rather on the sexual explicitness and prurience of the expression. *See Miller v. California*, 413 U.S. 15, 24-26 (1973).

12. *See* NADINE STROSSEN, *DEFENDING PORNOGRAPHY: FREE SPEECH, SEX, AND THE FIGHT FOR WOMEN’S RIGHTS* 114 (1995) (rejecting Catharine MacKinnon’s notion of false consciousness as paternalistic and subordinating of women).

13. For example, the marketplace rationale requires a genuine “consumer” of ideas such that the speech that prevails in the marketplace can be said to reflect the truest or most useful ideas; an assumption of false consciousness undermines the marketplace model in speech theory as in economic theory. *See Abrams v. United States*, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting). Likewise, the self-governance rationale of the First Amendment posits citizens capable of deliberative democratic choice. *See ALEXANDER MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT* (1948). And the personal autonomy rationale for protection of freedom of speech requires an individual able to exercise the desired autonomy. *See, e.g.*, FREDERICK SCHAUER, *FREE SPEECH: A PHILOSOPHICAL ENQUIRY* (1982).

vidual behavior is to some extent influenced by the larger social context in which it occurs.¹⁴ That larger context includes other people, institutional culture and architecture, social norms, legal rules, and physical features of the environment. Indeed, law would have little purpose absent the assumption that human behavior tends to mold to its shape.¹⁵ Humans are social animals; our behavior is impacted by the rules and other social cues present in our environment.

Decades of research in social psychology have demonstrated that individual behavior is strongly influenced by situation, and that disposition (an individual's "character" or "personality") does not have as strong an influence as people generally believe.¹⁶ In addition, numerous experiments performed over the past twenty years make clear that when certain social identity characteristics are "primed," individual behavior and decisionmaking are unconsciously affected in measurable ways.¹⁷ Finally, findings in developmental psychology and neuropsychology suggest that there may be good reason to treat college-age individuals differently than older adults.¹⁸

The exceedingly modest proposal¹⁹ of this Essay is that universities can, and should, consider ways in which to structure their social and physical environments so as to minimize harmful antisocial speech and to maximize prosocial, productive speech.²⁰ Rather than prohibit

14. The importance of situational factors to behavior is a commonplace belief among psychologists, though people in general tend to minimize the role of situation relative to disposition. See Adam Benforado & Jon Hanson, *The Great Attributional Divide: How Divergent Views of Human Behavior Are Shaping Legal Policy*, 57 EMORY L.J. 311, 321-22 (2008). To the extent that one believes that behavior is independent of situational factors, this Essay's proposal should not be objectionable, but merely ineffectual.

15. Cf. Owen D. Jones, *Time-Shifted Rationality and the Law of Law's Leverage: Behavioral Economics Meets Behavioral Biology*, 95 NW. U. L. REV. 1141, 1187 (2001) ("We can usefully consider law to be a lever for moving human behavior in directions it would not go on its own.").

16. See Jon Hanson & David Yosifon, *The Situation: An Introduction to the Situational Character, Critical Realism, Power Economics, and Deep Capture*, 152 U. PA. L. REV. 129, 165-66 (2003) (summarizing social psychological literature showing that situation, rather than disposition, is the crucial factor driving human behavior). Though well accepted in psychology literature, the importance of situation relative to disposition has been underappreciated in the legal academic literature. See Benforado & Hanson, *supra* note 14, at 315 (noting that "one purpose of the critical realist project . . . is to encourage and expedite th[e] process" of disseminating situationalist research among legal academics and policy-makers).

17. For a recent overview and synthesis of the research, see Ap Dijksterhuis et al., *Effects of Priming and Perception on Social Behavior and Goal Pursuit*, in SOCIAL PSYCHOLOGY AND THE UNCONSCIOUS: THE AUTOMATICITY OF HIGHER MENTAL PROCESSES 51 (John A. Bargh ed., 2007) [hereinafter Dijksterhuis et al., *Effects of Priming*].

18. See Jennifer Ann Drobac, *"Developing Capacity": Adolescent "Consent" at Work, at Law, and in the Sciences of the Mind*, 10 U.C. DAVIS J. JUV. L. & POL'Y 1, 19 (2006) (summarizing neuroscientific findings on adolescent brain development and noting that "[t]his new research confirms that adolescent brain development extends into the twenties").

19. Cf. Strossen, *supra* note 2, at 484.

20. Such features would also be likely to influence nonspeech behavior in a similar direction. In addition, most of the general ideas presented here—though not the specific details—could apply in other institutional contexts such as workplaces or prisons.

hate speech or punish it after it occurs,²¹ universities should take steps to create and foster a social and physical atmosphere in which it becomes less likely that hate speech will occur in the first place.²² Such an approach would recognize the social constructionist insight that belief and behavior are profoundly influenced—often unconsciously—by cultural practices, language,²³ and images. At the same time, it would avoid the heavy-handed censorship which is most objectionable to civil libertarians.²⁴

Following this Introduction, Part II sets out the distinctive features of the university environment that make it a unique—and potentially very fruitful—forum in which to experiment with strategies to encourage prosocial behavior and thereby reduce antisocial speech, including hate speech. Part III describes recent work in the field of so-

21. I do not specifically address in this Essay hate crimes or hate speech that the First Amendment would not protect, such as threats directed at individuals. Though the premise of the Essay is that these would also become less common and less violent and aggressive in a more prosocial environment, those unprotected acts or speech that still occur could—and should—be punished under established First Amendment doctrine. *See, e.g.*, *Virginia v. Black*, 538 U.S. 343, 360 (2003) (stating that threats of violence directed at a particular person or group and intended to “plac[e] the victim in fear of bodily harm or death” may be punished as unprotected speech); *Wisconsin v. Mitchell*, 508 U.S. 476, 490 (1993) (holding that a sentence enhancement for bias crimes did not violate the Defendant’s free speech rights); *R.A.V. v. City of St. Paul*, 505 U.S. 377, 389 (1992) (suggesting that sexually harassing speech may constitutionally be regulated because it “violate[s] laws directed not against speech but against conduct”). I should also point out that there are additional ways an institution might deal with hate speech incidents besides prohibition and punishment. As Professor Calleros has argued, an incident can serve as the impetus for education, advocacy, and empowerment even if—or perhaps because—it is not subject to formal disciplinary procedures. *See* Charles R. Calleros, *Paternalism, Counterspeech, and Campus Hate-Speech Codes: A Reply to Delgado and Yun*, 27 ARIZ. ST. L.J. 1249, 1256-63 (1995).

22. This was the strategy recommended by a commission at the University of Pennsylvania that was charged with examining the University’s policies in the wake of the so-called “water buffalo” incident in the early 1990’s. *See* COMM’N ON STRENGTHENING THE CMTY., UNIV. OF PA., FINAL REPORT: COMMISSION ON STRENGTHENING THE COMMUNITY (1994), http://www.upenn.edu/almanac/issues/past/strength_community_report.html [hereinafter FINAL REPORT].

23. There is an ongoing debate among linguists and social theorists concerning the degree to which language influences thought, or vice versa. *See* Seaman, *supra* note 9, at 412 n.288. Recent studies seem to support the “Whorfian” position that language does have a role in shaping perception. *See* Christine Kenneally, *When Language Can Hold the Answer*, N.Y. TIMES, Apr. 22, 2008, at F1 (citing and describing several recent studies).

24. In a recent article, two scholars who have been prominent in bringing the situational model of human behavior to a legal academic audience noted the close correlations between political conservatism and dispositionism on the one hand and political liberalism and situationism on the other. *See* Benforado & Hanson, *supra* note 14, at 382-99. Generally speaking, political conservatives tend to attribute outcomes to disposition (personality), whereas political liberals tend to attribute outcomes to situation (outside forces). *See id.* In line with these observations, it would be expected that strong free speech proponents would tend to have a dispositional view of human behavior while proponents of hate speech regulation would tend to have a more situational view. Rather than enter the attitudinal fray, I prefer to emphasize that if the situational view were to turn out to be incorrect or misguided, the policy proposals suggested here would, for the most part, simply be ineffectual and not pernicious. With that said, much research demonstrates that subtle situational features can have significant effects on behavior. *See* discussion *infra* Part III.A.

cial psychology that supports the view that individual behavior responds strongly to situational cues. In particular, this Part focuses on the deindividuation and priming literatures to demonstrate that subtle changes in environment can result in significant shifts in thought and behavior. In Part IV, the Essay concludes by offering some preliminary ideas as to what steps institutions might explore as a way to shape their environments to accomplish these goals.

II. HATE SPEECH AND THE UNIVERSITY

The university environment is distinctive for a number of reasons. First, only public institutions are technically bound to comply with the dictates of the First Amendment, even though many private institutions profess a commitment to adhere to its principles.²⁵ This allows private colleges and universities some degree of leeway in choosing how best to respond to hate speech. Second, the principle of academic freedom may alter the First Amendment analysis that applies to public universities as compared to other state actors.²⁶ Third, from the perspective of developmental psychology, the typical university student is not a child, nor is he or she²⁷ fully an adult. There is substantial evidence that the average human brain does not reach its full

25. Note that California schools are bound by state law to adhere to the First Amendment, and Stanford's hate speech code was struck down as unconstitutional under this state scheme. See *Corry v. Leland Stanford Junior Univ.*, No. 740309 (Cal. Super. Ct. Feb. 27, 1995) (holding that California's Leonard Law, Education Code section 94367, permitted plaintiffs to challenge free speech violation by private university notwithstanding the state action doctrine), available at <http://www.ithaca.edu/faculty/eduncan/265/corryvstanford.htm>.

26. A glance through the many cases collected and reported on the website of the Foundation for Individual Rights in Education, <http://www.thefire.org>, reveals that classroom speech by faculty members is often a target of enforcement under hate speech codes. In such situations, the crosscurrents of free speech and academic freedom protections result in a complicated puzzle. See Frederick Schauer, *Is There a Right to Academic Freedom?*, 77 U. COLO. L. REV. 907, 908 (2006) ("Although it is surely correct that there are robust constitutional rights to freedom of speech, freedom of association, and the free exercise of religion, to take just a small number of pertinent examples, the doctrinal, conceptual, and normative issues surrounding the idea of academic freedom are far murkier."). See generally J. Peter Byrne, *Constitutional Academic Freedom After Grutter: Getting Real About the "Four Freedoms" of a University*, 77 U. COLO. L. REV. 929 (2006) (describing development and contours of constitutional academic freedom doctrine and evaluating the impact of the Supreme Court's decision in *Grutter v. Bollinger*, 539 U.S. 306 (2003), upon the doctrine); Smolla, *supra* note 1, at 216-24 (discussing concept of academic freedom in the context of campus hate speech regulation and arguing that a public university constitutionally could proscribe a professor's racist, sexist, or homophobic hate speech in the classroom but not in the open forums on the campus); R. George Wright, *The Emergence of First Amendment Academic Freedom*, 85 NEB. L. REV. 793, 793 (2007) (noting that "[t]he idea of a constitutionally protected realm of academic freedom is controversial and judicially unsettled" and proposing that state restrictions of academic freedom generally be subject to strict scrutiny).

27. There is some evidence that the pronoun matters in this context; development is not, on average, identically timed for males and females. See Evelyn K. Lambe, *Dyslexia, Gender, and Brain Imaging*, 37 NEUROPSYCHOLOGIA 521 (1999); Deborah P. Waber, *Sex Differences in Cognition: A Function of Maturation Rate?*, 192 SCIENCE 572 (1976). The gender implications of the argument, however, are well beyond the scope of this Essay.

moral or cognitive maturity until individuals are in their mid-twenties.²⁸ To the extent that most college students are in their late teens and early twenties, some arguments on either side of the hate speech debate might apply differently to them. Fourth, because of the relatively contained nature of university life, modifications of the environment to encourage prosocial behavior are more practicable than they would be, for example, on the Internet or in the public square.²⁹ And finally, the special mission of the university might in some cases justify institutional behavior that would be less supportable outside the collegiate gates.³⁰ For these reasons, this Essay focuses on the university as a somewhat special case in considering the arguments for and against hate speech codes and in considering the actions a university might take in an attempt to minimize hate speech on campus.

Public university hate speech restrictions that have been challenged in court have been decisively struck down on First Amendment grounds.³¹ In addition, as noted above, most private institutions insist rather vociferously that they are committed to upholding freedom of speech on campus and that they will therefore follow the dictates of the First Amendment although they are not legally bound to do so.³² However, a comprehensive empirical examination by Professor Jon Gould of the practices of both public and private institutions, performed after the federal court decisions striking down campus hate speech regulations, has revealed that a large percentage of institutions adopt strategies of noncompliance, either passive or active.³³ In exploring the paradox of institutions' stated intent to adhere to First Amendment doctrine contrasted with actions that had precisely the

28. See *infra* notes 50-56 and accompanying text.

29. It may be more difficult (practically, if not legally) for public universities to implement structural or other changes to affect the nature of the speech of members of the general public who come onto campus to distribute literature or otherwise participate in debate. Recently, for example, Emory University decided to restrict admission to a talk given by a controversial speaker because of disruption at an earlier event by individuals not affiliated with the university. See Salvador Rizzo, *Outside Group Stifles Horowitz Speech*, EMORYWHEEL.COM, Oct. 25, 2007, <http://www.emorywheel.com/detail.php?n=24510>.

30. Cf. *Grutter v. Bollinger*, 16 F. Supp. 2d 797 (E.D. Mich. 1998).

31. See *Bair v. Shippensburg Univ.*, 280 F. Supp. 2d 357, 372 (M.D. Pa. 2003); *Dambrot v. Cent. Mich. Univ.*, 839 F. Supp. 477, 482 (E.D. Mich. 1993); *UWM Post v. Bd. of Regents*, 774 F. Supp. 1163, 1177 (E.D. Wis. 1991); *Doe v. Univ. of Mich.*, 721 F. Supp. 852, 866-67 (E.D. Mich. 1989); see also *Corry v. Leland Stanford Junior Univ.*, No. 740309 (Cal. Super. Ct. Feb. 27, 1995), available at <http://www.ithaca.edu/faculty/cduncan/265/corryvstanford.htm>.

32. See Gould, *College Hate Speech Codes*, *supra* note 7; Smolla, *supra* note 1. Note, however, that there are arguments that private universities may be contractually bound based upon statements in student handbooks or other literature. See FIRE – Foundation for Individual Rights in Education, <http://www.thefire.org> (last visited Nov. 12, 2008).

33. Gould distinguished between those institutions that left existing hate speech restrictions in place—which he characterized as passive noncompliers—and those which actually enacted new or more restrictive codes following the court decisions—which he characterized as active noncompliers. See Gould, *College Hate Speech Codes*, *supra* note 7, at 365.

opposite effect, Professor Gould discovered that the reasons that universities kept or created hate speech codes were multifaceted. In some cases, administrators decided that eliminating even an unconstitutional code would send a signal to minorities on campus that they were not valued members of the academic community.³⁴ In addition, several university leaders suggested in interviews that they had weighed the likely public relations effects of eliminating the codes against an unstated policy of nonenforcement.³⁵ In other words, administrators sometimes chose to keep the codes in place as a symbol of their commitment to equality and nondiscrimination, while at the same time pursuing an unwritten policy of underenforcement or nonenforcement so as not to make First Amendment waves in the media.³⁶

Administrators, it seems, are keenly aware of the symbolic value of having in place strong policies against racist (and other forms of) hate speech. As Professor Mari Matsuda argued in an influential early article, “[i]n a society that expresses its moral judgments through the law, and in which the rule of law and the use of law are characteristic responses to many social phenomena, [the] absence of laws against racist speech is telling.”³⁷ Critical race theorists describe at least two distinct harms of hate speech: the harm that flows from the hate speech itself and the second-order harm that results from official tolerance of the speech.³⁸ Hate speech codes attempt to reduce the first

34. See *id.* at 371-75; see also *Univ. of Mich.*, 721 F. Supp. at 864-65.

35. See Gould, *College Hate Speech Codes*, *supra* note 7, at 371-75.

36. See *id.* Universities are apt to receive unwelcome media attention when an apparent hate speech incident occurs, no matter what response they pursue. For example, in the aftermath of the “water buffalo” incident at the University of Pennsylvania, the University was condemned for being the site of the speech and then again for its handling of the disciplinary proceeding against the student who uttered the remark. See Elizabeth Rossi, *Aftermath, Consequences of Decade-Old Water Buffalo Affair*, DAILY PENNSYLVANIAN, Apr. 18, 2003 (stating that the handling of this and a related incident “brought national attention to Penn, and then-University President Sheldon Hackney and other administrators were accused of what many critics called political correctness run amuck”). The social consequences in terms of intergroup relations of the media coverage and the ensuing debate are arguably worse than those of the incident itself. *But see* Calleros, *supra* note 21, at 1256-63 (arguing that racist incidents on campus can serve as a “wake-up call” that may ultimately lead to empowerment, dialogue, and mutual understanding).

37. Matsuda, *supra* note 1, at 2379; see also Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021 (1996). Similar to feminist arguments for rape law reform in the 1970s, Professor Matsuda argued that “the law’s failure to provide recourse to persons who are demeaned by the hate messages is an effective second injury to that person . . . the pain of knowing that the government provides no remedy, and offers no recognition of the dehumanizing experience that victims of hate propaganda are subjected to.” Matsuda, *supra* note 1, at 2379. In an addendum to its decision in *University of Michigan*, the district court noted that it had become aware of the Matsuda article, delivered to chambers the very day that the decision was issued, just after its original opinion had been filed. *Univ. of Mich.*, 721 F. Supp. at 869. The court praised Matsuda’s article and stated that “[a]n earlier awareness of Professor Matsuda’s paper certainly would have sharpened the Court’s view of the issues.” *Id.*

38. According to Professor Matsuda:

kind of harm through deterrence and the second by their very existence and enforcement. Indeed, some scholars have argued that this second function is more important. Henry Louis Gates, Jr. has written that

the main appeal of speech codes usually turns out to be primarily expressive or symbolic rather than consequential in nature. That is, their advocates do not depend on the claim that the statute will spare victim groups some foreseeable amount of psychic trauma. They say, rather, that by adopting such a statute, the university *expresses* its opposition to hate speech and bigotry.³⁹

Among the most powerful arguments in favor of hate speech restrictions are those that stress that the harms caused by racist, sexist, and antigay hate speech are genuine, severe, and not recognized or considered sufficiently weighty by American law and society.⁴⁰ As Matsuda points out, “[t]olerance of hate speech is not tolerance borne by the community at large. Rather, it is a psychic tax imposed on those least able to pay.”⁴¹ Professor Richard Delgado has also catalogued at length the physical and psychological harms suffered by victims of racist hate speech and has argued that these harms can have perva-

To be hated, despised, and alone is the ultimate fear of all human beings. However irrational racist speech may be, it hits right at the emotional place where we feel the most pain. The aloneness comes not only from the hate message itself, but also from the government response of tolerance.

Matsuda, *supra* note 1, at 2338.

39. Henry Louis Gates, Jr., *War of Words: Critical Race Theory and the First Amendment*, in *SPEAKING OF RACE, SPEAKING OF SEX: HATE SPEECH, CIVIL RIGHTS, AND CIVIL LIBERTIES* 17, 38 (1994) (arguing that hate speech codes are misguided); *see also Univ. of Mich.*, 721 F. Supp. at 855 (quoting Acting University President’s statement that adoption of a hate speech policy “‘would enable the University to take the position that it was willing to do something about this issue [of racial harassment on campus]’”). On the other hand, as noted above, universities also wish to express their commitment to academic freedom and freedom of speech. A view of hate speech codes that understands them primarily as expressive or symbolic in nature raises complicated issues of academic freedom of the institution and, in the case of public universities, of the permissible scope of government speech. Of course, an institution could express its abhorrence of racist and sexist hate speech without purporting to restrict or punish such speech.

40. *See, e.g.*, DELGADO & STEFANCIC, *supra* note 5, at 11-18; ALEXANDER TESISIS, *DESTRUCTIVE MESSAGES: HOW HATE SPEECH PAVES THE WAY FOR HARMFUL SOCIAL MOVEMENTS* (2002); Delgado, *Words That Wound*, *supra* note 3, at 135-49; Lawrence, *supra* note 5, at 457-76; Matsuda, *supra* note 1, at 2326-41; Nielsen, *supra* note 3 (study measuring the nature, amount, and harms of racist and sexist remarks in public); *cf.* LEE C. BOLLINGER, *THE TOLERANT SOCIETY* 18-21 (1986) (describing Wigmore’s reaction to Justice Holmes’s famous dissent in *Abrams v. United States* and in particular Wigmore’s concern that in asserting a strong free speech principle protecting extremist speech Holmes “had grossly minimized the risk to the country from the defendants’ speech”).

41. Matsuda, *supra* note 1, at 2323; *see also* Nielsen, *supra* note 3, at 279 (“There can be little doubt that members of traditionally disadvantaged groups face a strikingly different reality on the street than do members of privileged groups.”).

sive and long-term effects upon the person's well-being.⁴² To the extent that institutions of higher education have a special responsibility to protect the physical, emotional, and psychological health of their students, the existence of these potential harms—and a recognition by the university that they are real and are serious—becomes increasingly relevant to the hate speech debate.

In addition to this focus on the victims of hate speech, the university context affords a special opportunity to examine the speaker side of the hate speech equation. While context is undoubtedly relevant in evaluating the listener side⁴³—and the university context arguably would support greater restrictions on speech for the purpose of ensuring equality of educational opportunity—on the speaker side, a university offers the special case in which speakers—whether fellow students or faculty—also have a particular demographic and a distinctive institutional role. Under these circumstances, the familiar First Amendment arguments might play out in a slightly different way.

It is well-settled that children do not enjoy the same level of constitutional protection of their freedom of speech as do adults.⁴⁴ In the educational context, though students do not “shed their constitutional rights to freedom of speech or expression at the schoolhouse gate,”⁴⁵ their interests in expression⁴⁶ must be balanced against the interests

42. Professor Delgado argues that “[t]he person who is timid, withdrawn, bitter, hypertense, or psychotic [as a consequence of being subjected to hate speech] will almost certainly fare poorly in employment settings.” Delgado, *Words That Wound*, *supra* note 3, at 139. In response to Delgado's argument, Professor Gates quipped, “As a member of the Harvard faculty, I would venture there are exceptions to this rule.” Gates, *supra* note 39, at 24.

43. For example, speech restrictions in the workplace have been permitted in order to ensure equal employment opportunity under Title VII, though it is difficult in principle to distinguish them from similar speech restrictions in other contexts. See *R.A.V. v. City of St. Paul*, 505 U.S. 377 (1992); Kingsley R. Browne, *Zero Tolerance for the First Amendment: Title VII's Regulation of Employee Speech*, 27 OHIO N.U. L. REV. 563 (2001); Kingsley R. Browne, *Title VII as Censorship: Hostile-Environment Harassment and the First Amendment*, 52 OHIO ST. L.J. 481 (1991); Eugene Volokh, *Deterring Speech: When is it “McCarthyism”? When is it Proper?*, 93 CAL. L. REV. 1413, 1416 (2005) [hereinafter Volokh, *Deterring Speech*]; Eugene Volokh, *What Speech Does “Hostile Work Environment” Harassment Law Restrict?*, 85 GEO. L.J. 627 (1997); Eugene Volokh, *Freedom of Speech and Appellate Review in Workplace Harassment Cases*, 90 NW. U. L. REV. 1009 (1996).

44. *E.g.*, *Morse v. Frederick*, 127 S. Ct. 2618, 2622 (2007) (“[W]e have held that ‘the constitutional rights of students in public school are not automatically coextensive with the rights of adults in other settings,’ and that the rights of students ‘must be applied in light of the special characteristics of the school environment.’” (citation omitted)); *FCC v. Pacifica Found.*, 438 U.S. 726, 757 (1978) (noting that it is permissible to restrict children's access to speech that would be protected as to adults because “a child . . . is not possessed of that full capacity for individual choice which is the presupposition of First Amendment guarantees” (quoting *Ginsberg v. New York*, 390 U.S. 629, 649-50 (1968) (Stewart, J., concurring))).

45. *Tinker v. Des Moines Indep. Sch. Dist.*, 393 U.S. 503, 506 (1969).

46. Students' privacy interests are similarly constrained relative to adult citizens. See *Bd. of Educ. of Indep. Sch. Dist. No. 92 of Pottawatomie County v. Earls*, 536 U.S. 822 (2002) (holding that policy requiring all students who participated in competitive extracurricular activities to submit to suspicionless drug testing did not violate Fourth Amendment ban on unreasonable searches); *Vernonia Sch. Dist. 47J v. Acton*, 515 U.S. 646, 657 (1995)

of the state acting as *parens patriae*⁴⁷ and the interests of other students in the educational environment.⁴⁸ University students, however, have traditionally been treated as adults for purposes of First Amendment analysis.⁴⁹

A growing body of research suggests, however, that significant brain development continues into the early- and mid-twenties—that is, into the typical college years of most undergraduate and even graduate students. “The relatively recent discovery that brain development continues into adulthood has turned the conventional wisdom that the brain is fully developed by early childhood on its head.”⁵⁰ This relatively late brain development is concentrated in areas of the brain that involve executive function, including moral decisionmaking, impulse inhibition, and reasoning about the consequences of one’s actions. As summed up by one legal scholar:

[W]ork by such researchers as Piaget and (more so) Kohlberg showed a process of development in moral reasoning going far beyond early childhood. Neuroscience now provides an explanation of that extended period of development. While the brain may be sufficiently developed for many tasks in early childhood, when it comes to inhibition and understanding consequences vital to moral reasoning, the wiring for the task is not complete.⁵¹

Through a developmental process that has in recent years been revealed by new techniques of brain imaging,⁵² “the teenage brain fine-

(“ [S]tudents within the school environment have a lesser expectation of privacy than members of the population generally.” (quoting *New Jersey v. T.L.O.*, 469 U.S. 325, 348 (1985))).

47. *Vernonia*, 515 U.S. at 654. The Court stated:

Traditionally at common law, and still today, unemancipated minors lack some of the most fundamental rights of self-determination—including even the right of liberty in its narrow sense, *i.e.*, the right to come and go at will. They are subject, even as to their physical freedom, to the control of their parents or guardians.

Id. Public schools, according to the Court, both stand *in loco parentis* to minor children as well as exercise the power of the state. *Id.* at 654-55.

48. *Tinker*, 393 U.S. at 509 (stating that a school may restrict student speech when it has reason to believe that such expression will “impinge upon the rights of other students”).

49. According to Professor Eugene Volokh, “[t]he government acting as college educator is generally assumed by recent lower court cases to have no greater powers than the government acting as sovereign,” though “there has been no square Supreme Court holding” on the question. EUGENE VOLOKH, *THE FIRST AMENDMENT AND RELATED STATUTES* 379 (3d ed. 2008).

50. Kenneth J. King, *Waiving Childhood Goodbye: How Juvenile Courts Fail to Protect Children from Unknowing, Unintelligent, and Involuntary Waivers of Miranda Rights*, 2006 WIS. L. REV. 431, 437 n.24.

51. Kevin W. Saunders, *A Disconnect Between Law and Neuroscience: Modern Brain Science, Media Influences, and Juvenile Justice*, 2005 UTAH L. REV. 695, 712 (citation omitted).

52. The development of functional magnetic resonance imaging (fMRI) has revolutionized brain science in recent years, allowing researchers to observe subjects’ brains as they perform cognitive and motor tasks and to see which areas of the brain are differentially

tunes its most human part, the prefrontal cortex, the place that helps us cast a wary eye, link cause to effect, decide ‘maybe not’—the part, in fact, that acts grown-up.’ The process takes place not only in the teen years, but into the twenties.”⁵³

Furthermore, adolescents in the midst of this rapid prefrontal development are also more susceptible than fully mature adults to cues from social groups with which they self-identify.⁵⁴ Individuals who are in the midst of this late teen/early adult developmental phase “are less self-reliant and more influenced . . . by others.”⁵⁵ Because of this heightened sensitivity to the judgments and values of the social groups with which they identify, it is even more likely that situational cues will have an effect on students’ behavior.⁵⁶

In sum, because of their neural, social, cognitive, educational, and moral developmental stage, college students as a group are probably more likely than older adults to engage in antisocial behavior where situational factors lead them in that direction, but they are also more likely to respond to situational and social cues that would tend to lead them instead toward prosocial behavior. From the perspective of the First Amendment, it is defensible to argue that postsecondary student speech, at least within the university context, should be subject to somewhat more regulation than the speech of fully mature adults. Thus, universities should be permitted some room to experiment with situational manipulation in order to encourage prosocial behavior and to discourage antisocial behavior and speech, including hate speech. The following Part describes the social science research that suggests why situational manipulation of this sort might be successful.

“working.” See Jay N. Geidd et al., *Brain Development During Childhood and Adolescence: A Longitudinal MRI Study*, 2 NATURE NEUROSCIENCE 861 (1999); Tomáš Paus, *Mapping Brain Maturation and Cognitive Development During Adolescence*, 9 TRENDS COGNITIVE SCI. 60 (2005); Julie A. Seaman, *Black Boxes*, 58 EMORY L.J. (forthcoming 2009) (detailing recent advances in functional Magnetic Resonance Imaging (fMRI)).

53. Saunders, *supra* note 51, at 710 (quoting BARBARA STRAUCH, THE PRIMAL TEEN: WHAT THE NEW DISCOVERIES ABOUT THE TEENAGE BRAIN TELL US ABOUT OUR KIDS 203-04 (2003)).

54. In other words, they are more susceptible to peer pressure. See Drobac, *supra* note 18, at 28-29; Cass R. Sunstein, *Adolescent Risk-Taking and Social Meaning: A Commentary* 2-3 (John M. Olin Law & Econ. Working Paper No. 386 (2d series), 2008, Public Law and Legal Theory Working Paper No. 198, 2008), available at http://ssrn.com/abstract_id=1087144.

55. Drobac, *supra* note 18, at 29.

56. Much research in the fields of social and cognitive psychology is done using students—often undergraduates—as subjects. This common feature of the experimental design is sometimes criticized for the very reason that it is not clear how results will generalize outside of this population. However, for the purposes of the analysis presented in this Essay, it is especially felicitous.

III. SOCIAL PSYCHOLOGY: INDIVIDUALS, GROUPS, AND BEHAVIOR

The field of social psychology offers a host of insights into human attitudes, perceptions, and behaviors. Many studies performed in recent years support the proposition that situational elements—the physical, institutional, social, cultural and linguistic environment in which a person is situated—affect individual thought and action in measurable and fairly predictable ways. In particular, this Part describes research in the areas of deindividuation and priming, and it proposes that these two literatures complement one another and suggest potentially fruitful avenues by which universities could encourage prosocial speech and behavior.

A. *Deindividuation Theory and the Psychology of Anonymity*

Psychologists (and others) have long been interested in the question of what drives individuals to behave in either prosocial or antisocial ways. More specifically, many social commentators have noticed that individuals in groups often behave differently—and worse—than individuals acting alone.⁵⁷ Beginning with the 1952 publication of a classic paper on the subject, the predominant framework for understanding so-called mob behavior has been the theory of “deindividuation.”⁵⁸ Social psychologist Leon Festinger⁵⁹ hypothesized that when individuals are “submerged in the group,” they experience a relaxation of inhibition and social restraint and are more apt to act antisocially.⁶⁰ Somewhat paradoxically, according to this theory, individuals in groups *feel* more anonymous than when they are alone; that feeling of relative anonymity—or deindividuation—leads to antinormative behavior.⁶¹

In the years since Festinger’s article was published, numerous studies have sought to explain the behavior of individuals within groups and to identify the factors that contribute to socially destruc-

57. See Tom Postmes, *Deindividuation*, in 1 ENCYCLOPEDIA OF SOCIAL PSYCHOLOGY 233 (Roy F. Baumeister & Kathleen D. Vohs eds., 2007) (tracing deindividuation theory to the classic 1895 book *La Foule* (The Crowd), by Gustave Le Bon); see also NATHANAEL WEST, *THE DAY OF THE LOCUST* (1950).

58. See L. Festinger et al., *Some Consequences of De-Individuation in a Group*, 47 J. ABNORMAL & SOC. PSYCHOL. 382, 382 (1952).

59. Festinger is especially well known for introducing the theory of cognitive dissonance. See LEON FESTINGER, *A THEORY OF COGNITIVE DISSONANCE* (1957).

60. Festinger et al., *supra* note 58, at 382.

61. The social psychology literature refers to aggressive, dishonest, or otherwise undesirable behavior as antinormative, in that it contravenes generally accepted social norms. However, this label is problematic; more recent research suggests that the behavior may in fact be normative with respect to the smaller group, though it is antinormative with respect to the larger group. See *infra* notes 71-77 and accompanying text.

tive dynamics.⁶² Though there has at times been disagreement about the explanation for observed correlations between certain situational features and antisocial behaviors, there is little dispute that the correlations exist.⁶³ The most persistent connection across many studies has been between anonymity of various kinds or degrees and changes in behavior. Thus, when researchers manipulate the extent to which subjects perceive themselves as individuals versus as members of groups, the behavior of the subjects change, sometimes dramatically so.

In *The Lucifer Effect*, social psychologist Philip Zimbardo examined forty years of empirical research to conclude that context is critically important in shaping behavior.⁶⁴ Decades of research have demonstrated that certain situational factors strongly predict the likelihood of antinormative behavior. In particular, social psychologists have long recognized the critical importance of identifiability versus anonymity as a factor influencing behavior.⁶⁵ Countless laboratory and field studies, beginning with Zimbardo's infamous Stanford Prison Experiment in 1971 and continuing to the present with studies of internet speech and behavior, have demonstrated that a feeling of anonymity can often increase—even cause—aggressive, disturbing, and antisocial behavior, including speech.⁶⁶

62. For purposes of this discussion, I make what I hope is the uncontroversial assumption that aggression, violence, sadism, degradation of another person, racism, and behavior that generally falls under the rubric of “hate speech” is antisocial and undesirable.

63. *But see* Tom Postmes & Russell Spears, *Deindividuation and Antinormative Behavior: A Meta-Analysis*, 123 PSYCHOL. BULL. 238 (1998) (noting the anomalies in the research and proposing refinements to the theory) [hereinafter Postmes & Spears, *A Meta-Analysis*].

64. *See* PHILIP ZIMBARDO, *THE LUCIFER EFFECT: UNDERSTANDING HOW GOOD PEOPLE TURN EVIL* (2007).

65. *See* Katelyn Y. A. McKenna & John A. Bargh, *Plan 9 From Cyberspace: The Implications of the Internet for Personality and Social Psychology*, 4 PERSONALITY & SOC. PSYCHOL. REV. 57, 61-62 (2000) (summarizing literature on deindividuation and stating that deindividuation, “as through anonymity . . . decreases the influence of internal (i.e., self) standards of or guides to behavior, and increases the power of external, situational cues”).

66. The distinction between “speech” (for First Amendment purposes) and conduct is notoriously thorny. Some nonverbal conduct—such as flag-burning—is protected as “symbolic speech,” while some verbal or written utterances—such as threats, bribes, or revealing state secrets—are often said to lack First Amendment protection because they are not really speech but instead are akin to conduct. *See generally* Eugene Volokh, *Speech as Conduct: Generally Applicable Laws, Illegal Courses of Conduct, “Situation-Altering Utterances,” and the Uncharted Zones*, 90 CORNELL L. REV. 1277 (2005). The debate over the constitutionality and desirability of regulating hate speech has highlighted this issue, with proponents of hate speech regulation often likening hate speech to physical assault. *See, e.g.*, Richard Delgado, *Are Hate-Speech Rules Constitutional Heresy? A Reply to Steven Gey*, 146 U. PA. L. REV. 865, 872 (1998) (“Face-to-face hate speech conveys no information. It is more like a slap in the face or a performative . . .”); Matsuda, *supra* note 1, at 2332 (“In addition to physical violence, there is the violence of the word. Racist hate messages, threats, slurs, epithets, and disparagement all hit the gut of those in the target group.”). Putting aside this debate under the First Amendment, however, for purposes of this Essay the operative point is simply that within social psychology, speech is treated as a type of behavior.

In the Stanford Prison Experiment, subjects who were by all accounts average, well-adjusted, and well-educated⁶⁷ young men were randomly assigned to play the roles either of prisoners or of prison guards for a two-week study of the psychology of imprisonment. Within the first week, the study was called off as the behavior of the “guards” toward the “prisoners” became increasingly abusive, aggressive, violent, and sadistic and the “prisoners” began to experience severe emotional distress.⁶⁸ Notably, certain features emerged within the social setting of the experiment that have since been shown to correlate with similarly aggressive, dehumanizing behavior toward other individuals. In particular, social cues that tended to minimize the individuality of the targets of the behavior—such as numbers replacing names, standard prison uniforms replacing individual items of clothing, and covering of the victims’ faces—were implemented by the “guards” at Stanford.⁶⁹ Likewise, features were adopted which tended to minimize the individuality of the guards as well. These included standard military-style uniforms and silver reflecting sunglasses.⁷⁰

This experiment, along with a substantial body of research on deindividuation,⁷¹ has since been widely interpreted as suggesting that “anything, or any situation, that makes people feel anonymous, as though no one knows who they are or cares to know, reduces their sense of personal accountability, thereby creating the potential for evil action.”⁷² From this perspective, the issue is not whether to tolerate

67. See Craig Haney et al., *Interpersonal Dynamics in a Simulated Prison*, 1 INT’L J. CRIMINOLOGY & PENOLOGY 69, 73 (1973); ZIMBARDO, *supra* note 64, at 32-33.

68. See ZIMBARDO, *supra* note 64, at 171.

69. See *id.* at 40. Zimbardo notes that, from the very beginning of the experiment and “[w]ithout any staff encouragement, some guards begin to make fun of the prisoners’ genitals, remarking on their small penis size or laughing at their unevenly hanging testicles.” *Id.*

70. *Id.* at 301.

71. “Theories of deindividuation propose that it is a psychological state of decreased self-evaluation and decreased evaluation apprehension causing antinormative and disinhibited behavior.” Postmes & Spears, *A Meta-Analysis*, *supra* note 63, at 238 (citations omitted). The theory developed from an attempt to explain observed mob behavior. See Tom Postmes, *Deindividuation*, in 1 ENCYCLOPEDIA OF SOCIAL PSYCHOLOGY, *supra* note 57, at 233 (“Deindividuation theory was developed to explain the violence and irrationality of the crowd. How does a group of seemingly normal individuals become an unruly mob? According to deindividuation theory, the anonymity and excitement of the crowd make individuals lose a sense of individual identity. As a result, crowd members cease to evaluate themselves, and they become irrational and irresponsible. All of this makes the crowd fickle, explosive, and prone to anti-normative and disinhibited behavior.”). Subsequent research has led some social psychologists to conclude that “[a]nonymity does not render people unthinkingly violent. Rather, anonymity increases their responsiveness to the normative cues present in their immediate environment.” *Id.* at 234.

72. ZIMBARDO, *supra* note 64, at 301; see also McKenna & Bargh, *supra* note 65, at 61 (citing Zimbardo study for the proposition that the effects of anonymity “can culminate in impulsive and disinhibited behaviors”); Michael J. White, *Counternormative Behavior as Influenced by Deindividuating Conditions and Reference Group Salience*, 103 J. SOC. PSYCHOL. 75, 76 (1977) (stating that “Zimbardo . . . demonstrated that anonymous college students were considerably more aggressive than those who were not anonymous”).

the few bad apples in order to preserve the liberty of the rest; rather, it is that certain (predictable) features of the barrel can cause even the good apples to go bad.⁷³

A more recent refinement of deindividuation theory argues that it is not necessarily (or only) the anonymity per se that leads to negative social behaviors; rather, the crucial causal factor is the set of social norms associated with the particular group or subgroup with which the (deindividuated) subject identifies. According to this theory—known as SIDE⁷⁴—an individual constantly identifies with himself or herself as well as with a range of social groups; where the norms of a subgroup support negative behavior,⁷⁵ the person who attaches salience to that subgroup identity will be more likely to behave in accordance with these negative social norms. The SIDE model

proposes that anonymity may induce a shift in focus from individual identity to social identity The principal reason for this is that the decreased visibility of individuality can shift the emphasis away from concerns about others' individuality within the group, and towards the shared communalities and group concerns.⁷⁶

This social identity perspective helps to explain why some individuals in groups do not exhibit classic negative mob behavior; at the same time, the social identity focus retains the explanatory value of the specific variables that can affect behavior (either for good or bad) by decreasing the sense of individuation.⁷⁷

73. “[T]he bad apple-dispositional view ignores the apple barrel and its potentially corrupting situational impact on those within it.” ZIMBARDO, *supra* note 64, at 10. Prior to the Stanford Prison Experiment, the prevailing wisdom regarding prison conditions was a dispositional view according to which, “despicable conditions, violence, brutality, dehumanisation and degradation existing within any prison can be traced to some innate or acquired characteristic of the correctional and inmate population.” Haney et al., *supra* note 67, at 70.

74. SIDE stands for “the social identity model of deindividuation effects.” For background and a general overview of the SIDE model and associate research, see Postmes & Spears, *A Meta-Analysis*, *supra* note 63, at 241-42.

75. As already noted, *see supra* note 61, it is somewhat misleading to label the behavior “antinormative,” since it is normative with respect to the subgroup. It is antinormative only with respect to some larger societal goal. In general, I assume for purposes of this discussion that it is both possible and appropriate to categorize some speech as undesirable, though it might be neither constitutional nor appropriate for the state to censor or punish it for that reason alone. *See* BOLLINGER, *supra* note 40, at 12 (noting the “curious disjunction” between Americans’ “readiness to employ a host of informal, or *nonlegal*, forms of coercion” against offensive speech and our attitude toward government censorship of the same speech); Eugene Volokh, *Deterring Speech*, *supra* note 43, at 1416 (“Yet as the example of social norms against racist speech shows, some deterrence of bad speech is socially and legally permissible.”).

76. Martin Tanis & Tom Postmes, *Two Faces of Anonymity: Paradoxical Effects of Cues to Identity in CMC*, 23 COMPUTERS HUM. BEHAV. 955, 959 (2007) (citations omitted). “CMC” refers to Computer Mediated Communications. *Id.* at 955.

77. Much of the recent work on deindividuation and anonymity has been done in the context of computer-mediated communication (CMC). *See* John A. Bargh & Katelyn Y. A. McKenna, *The Internet and Social Life*, 55 ANN. REV. PSYCHOL. 573 (2004); Tom Postmes et al., *Social Influence in Computer-Mediated Communication: The Effects of Anonymity on*

B. Social Identity and Priming

As the deindividuation research canvassed above suggests, both the level of group identification and the particular group with which a person identifies at a given time may influence behavior (including speech). Another body of experimental social psychology, which studies the so-called “priming” effect, also demonstrates the relevance of group identification and group trait attribution⁷⁸ to individual behavior as well as cognitive performance. Furthermore, as these priming studies make clear, the effects of social identity salience can operate outside of conscious awareness and can persist over time.⁷⁹

Group Behavior, 27 PERSONALITY & SOC. PSYCHOL. BULL. 1243 (2001); Tom Postmes & Russell Spears, *Behavior Online: Does Anonymous Computer Communication Reduce Gender Inequality?*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 1073 (2002) [hereinafter Postmes & Spears, *Behavior Online*]; Giuseppe Riva, *The Sociocognitive Psychology of Computer-Mediated Communication: The Present and Future of Technology-Based Interactions*, 5 CYBERPSYCHOLOGY & BEHAV. 581 (2002); Russell Spears et al., *When Are Net Effects Gross Products? The Power of Influence and the Influence of Power in Computer-Mediated Communication*, 58 J. SOC. ISSUES 91 (2002); John Suler, *The Online Disinhibition Effect*, 7 CYBERPSYCHOLOGY & BEHAV. 321 (2004); Tanis & Postmes, *supra* note 76. As noted above, deindividuation theory originated in an attempt to explain mob behavior. In that context, the idea was that persons feel more anonymous in groups and therefore act in ways that they would not if they were subject to individual scrutiny. Further research supported this idea, finding that factors that could be said to decrease individual salience led to increased antinormative behavior. See, e.g., ZIMBARDO, *supra* note 64, at 24-25 (summarizing research showing that “conditions that make us feel anonymous . . . can foster antisocial, self-interested behaviors”); Philip G. Zimbardo, *The Human Choice: Individuation, Reason, and Order Versus Deindividuation, Impulse, and Chaos*, in 1969 NEBRASKA SYMPOSIUM ON MOTIVATION 237, 263-70 (William J. Arnold & David Levine eds., 1970) (subjects who were deindividuated by wearing hoods and lab coats and by working in darkness were more likely to administer electric shocks to innocent victims); Eugene W. Mathes & Thomas A. Guest, *Anonymity and Group Antisocial Behavior*, 100 J. SOC. PSYCHOL. 257 (1976) (finding that both disguise and group membership facilitated feelings of anonymity and increased the antinormative behavior of willingness to carry a sign reading “masturbation is fun”). This focus on the anonymity aspect of the theory—rather than the social group aspect—made the model a natural fit in the Internet age of highly anonymous social interaction. Ironically, it is research in the Internet context that has led social psychologists back to a focus on the social group identity elements of deindividuation theory. See, e.g., Postmes et al., *supra*, at 1252 (finding that increased anonymity in computer-mediated communication can have the counterintuitive effect of enhancing group identification). The paradox is that a person sitting alone at a computer is highly influenced by the virtual social group with which he identifies precisely because he is increasingly anonymous. See, e.g., McKenna & Bargh, *supra* note 65; Postmes & Spears, *Behavior Online*, *supra*; Tanis & Postmes, *supra* note 76.

78. The priming effects discussed in this section can be invoked either by priming a trait directly (for example, “slow,” “smart,” “cooperative,” or “rude”) or by priming a social category that evokes the trait through attribution (for example, the category “elderly” evokes the trait “slow”). See Dijksterhuis et al., *Effects of Priming*, *supra* note 17, at 76. For an excellent discussion of the priming research and its legal relevance, see Jerry Kang, *Trojan Horses of Race*, 118 HARV. L. REV. 1489 (2005).

79. See generally SOCIAL PSYCHOLOGY AND THE UNCONSCIOUS: THE AUTOMATICITY OF HIGHER MENTAL PROCESSES (John A. Bargh ed., 2007). Regarding the priming effect’s persistence over time, see Leif D. Nelson & Michael I. Norton, *From Student to Superhero: Situational Primes Shape Future Helping*, 41 J. EXPERIMENTAL SOC. PSYCHOL. 423, 428-29 (2005).

The priming research reveals the effect upon behavior of very subtle—often subliminal—cues that invoke social categories or traits.⁸⁰ A common experimental design involves subjects participating in what they are told is a language study.⁸¹ The subjects are given a scrambled sentence or word association test in which are embedded certain words that evoke a social category.⁸² After the subject has finished the language task and is under the impression that the experiment is over, his or her conduct is observed.⁸³ For example, in an early experiment subjects were given language tasks that included words such as “wrinkled,” “cane,” and “gray.”⁸⁴ After they had completed the language task and believed that the experiment was completed, they walked significantly more slowly from the testing room to the elevator than control subjects who were not unconsciously primed for the social category “elderly.”⁸⁵ In similar experiments, behavioral effects have been observed based on primes for the social categories, including “business people” person (primed subjects moved more quickly),⁸⁶ “superhero” (primed subjects were more helpful),⁸⁷ “punk,” (primed subjects exhibited more nonconformity),⁸⁸ “professor” (primed subjects performed better on Trivial Pursuit questions),⁸⁹ “soccer hooligan” (primed subjects performed worse at Trivial Pursuit),⁹⁰ and “supermodel” (primed subjects performed less well on a test).⁹¹ It appears that people associate certain traits with particular social categories

80. See John A. Bargh et al., *Automaticity of Social Behavior: Direct Effects of Trait Construct and Stereotype Activation on Action*, 71 J. PERSONALITY & SOC. PSYCHOL. 230 (1996).

81. *Id.* at 233.

82. *Id.*

83. *Id.* at 234-35.

84. See *id.* at 236.

85. *Id.* at 237. These results have been replicated. See J. G. Hull et al., *The Nonconsciousness of Self-Consciousness*, 83 J. PERSONALITY & SOC. PSYCHOL. 406 (2002); K. Kawakami et al., *Automatic Stereotyping: Category, Trait, and Behavioral Activations*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 3 (2002).

86. Russell Spears et al., *Reaction in Action: Intergroup Contrast in Automatic Behavior*, 30 PERSONALITY & SOC. PSYCHOL. BULL. 605 (2004).

87. Nelson & Norton, *supra* note 79.

88. Louise Pendry & Rachael Carrick, *Doing What the Mob Do: Priming Effects on Conformity*, 31 EUR. J. SOC. PSYCHOL. 83 (2001).

89. See Ap Dijksterhuis & Ad van Knippenberg, *The Relation Between Perception and Behavior, or How to Win a Game of Trivial Pursuit*, 74 J. PERSONALITY & SOC. PSYCHOL. 865, 869-71 (1998).

90. See *id.* at 871-72.

91. See Ap Dijksterhuis et al., *Seeing One Thing and Doing Another: Contrast Effects in Automatic Behavior*, 75 J. PERSONALITY & SOC. PSYCHOL. 862 (1998) [hereinafter Dijksterhuis et al., *Seeing One Thing*]; Geoffrey Haddock et al., *Syrian Science and Smart Supermodels: On the When and How of Perception-Behavior Effects*, 20 SOC. COGNITION 461 (2002).

and that unconscious primes for the categories cause people to behave in accordance with the associated traits.⁹²

As might be expected, effects have also been observed in response to primes for gender, racial, and ethnic categories.⁹³ For example, participants in one study were asked to complete a particularly monotonous computer task, during which either White or African-American faces were flashed on the screen too rapidly to be accessible to the subjects' conscious awareness.⁹⁴ When the test was nearly complete, the computer appeared to crash, displaying an error message, and the subject was informed he or she would have to begin again.⁹⁵ Those individuals who had been subliminally exposed to an African-American face reacted with much more hostility than those exposed to a White face.⁹⁶ In another study, which demonstrated the effect of social identity priming upon the subject's intellectual performance, female Asian college math majors primed with the category "Asian" scored better on a math test, while those primed with the category "female" scored worse.⁹⁷ Control subjects primed with nonidentity words performed in between these groups.⁹⁸

92. Where subjects did not associate the category with the trait, the effect was not found. Bargh et al., *supra* note 80, at 240. Likewise, the stronger the association between the category and the trait, the more pronounced the behavioral effect. See Henk Aarts & Dijksterhuis, *The Silence of the Library: Environment, Situational Norm, and Social Behavior*, 84 J. PERSONALITY & SOC. PSYCHOL. 18, 26 (2003) (noting that research shows that "accessibility of concepts after priming depends on the associative strength between the concept and prime"). Though this effect is very robust, a few studies have shown a contrary effect in which the primed subject tends to behave opposite from the stereotypical trait. See *id.* at 86-91 (discussing "contrast effects"). Recently, scholars have developed a more refined model that explains the anomalies by showing that whether an assimilation or contrast effect is observed depends on whether, and how much, the primed identity is viewed as "in-group" or "out-group" by the subject. See Alison Ledgerwood & Shelly Chaiken, *Priming Us and Them: Automatic Assimilation and Contrast in Group Attitudes*, 93 J. PERSONALITY & SOC. PSYCHOL. 940, 940-41 (2007).

93. See Kai J. Jonas & Kai Sassenberg, *Knowing How to React: Automatic Response Priming from Social Categories*, 90 J. PERSONALITY & SOC. PSYCHOL. 709, 709 (2006) (stating that "[t]his phenomenon [whereby social categories activate attributes and behaviors stereotypical of the activated category], called *automatic stereotype activation*, has been demonstrated for numerous social categories such as African-Americans and other ethnics of color, women, the elderly, and gays").

94. Bargh et al., *supra* note 80, at 238-39.

95. *Id.* at 238.

96. *Id.* at 239; see also Mark Chen & John A. Bargh, *Nonconscious Behavioral Confirmation Processes: The Self-Fulfilling Consequences of Automatic Stereotype Activation*, 33 J. EXPERIMENTAL SOC. PSYCHOL. 541 (1997).

97. Margaret Shih et al., *Stereotype Susceptibility: Identity Salience and Shifts in Quantitative Performance*, 10 PSYCHOL. SCI. 80, 80-81 (1999).

98. See *id.* at 80-81. The phenomenon whereby activation of a "positive" stereotype improves performance is known in literature as "stereotype boost." See, e.g., Jessi L. Smith & Camille S. Johnson, *A Stereotype Boost or Choking Under Pressure? Positive Gender Stereotypes and Men Who Are Low in Domain Identification*, 28 BASIC & APPLIED SOC. PSYCHOL. 51 (2006). The corresponding negative affect is called "stereotype threat." See, e.g., Harriet E. S. Rosenthal et al., *Improving Performance Expectancies in Stereotypic Domains: Task Relevance and the Reduction of Stereotype Threat*, 37 EUR. J. SOC. PSYCHOL. 586 (2007).

The same effects have been observed where the trait is primed directly rather than through the mediator of the social category.⁹⁹ Thus, individuals primed with words associated with rudeness interrupted a confederate of the experimenter more quickly than those individuals primed with words associated with politeness, and control subjects primed with neither fell between these two groups.¹⁰⁰ In this manner, studies have shown behavioral effects for various trait primes, including cooperation, aggressiveness, helpfulness, intelligence, and stupidity.¹⁰¹ The priming literature demonstrates that individual and social behavior is amenable to being nudged in a prosocial direction by cues present in the environment.¹⁰² In particular, social interactions can be affected by subliminal or supraliminal primes¹⁰³ that invoke traits such as helpfulness, morality, and cooperation.¹⁰⁴

Interestingly, physical features of the environment can also influence behavior by invoking shared social norms regarding appropriate behavior within a given context. Referred to in the literature as “situational norms,” these consist of “generally accepted beliefs about how to behave in particular situations.”¹⁰⁵ Research has shown that normative behavior can be invoked by priming the physical cues associated with the particular behavior and that this can occur outside the conscious awareness of the individual.¹⁰⁶ In one study, subjects primed with a photograph of a library subsequently spoke more quietly when engaged in a supposedly unrelated experiment about language.¹⁰⁷ In another study, students who received handbills in their mailboxes threw them on the floor when the floor was already littered with handbills; when the floor was clean, they did not.¹⁰⁸ In addition, mun-

99. See Bargh et al., *supra* note 80, at 233-35.

100. *Id.*

101. See Betsy Sparrow & Daniel M. Wegner, *Unpriming: The Deactivation of Thoughts Through Expression*, 91 J. PERSONALITY & SOC. PSYCHOL. 1009, 1009-10 (2006) (summarizing research and citing sources).

102. See, e.g., Aarts & Dijksterhuis, *supra* note 92.

103. Subliminal primes (such as a very rapid flash of a word or photograph) are observed outside of conscious awareness; supraliminal primes (such as the words in a scrambled sentence task) are consciously observed, but the subject is unaware of the purpose or valence of the terms. See, e.g., Dirk Hermans et al., *Affective Priming with Subliminally Presented Pictures*, 57 CANADIAN J. EXPERIMENTAL PSYCHOL. 97 (2003); see also RANDOM HOUSE WEBSTERS COLLEGE DICTIONARY 1301, 1314 (2d ed. 1999) (defining “subliminal” in psychology as “existing or operating below the threshold of consciousness” and “supraliminal” as “being above the threshold of perception” of a stimulus).

104. See Dijksterhuis et al., *Effects of Priming*, *supra* note 17, at 82 (discussing the existing research).

105. Aarts & Dijksterhuis, *supra* 92, at 18.

106. See *id.* at 18-20, 25-27 (citing sources).

107. *Id.* at 21-22; see also Janneke F. Joly et al., *Silence and Table Manners: When Environments Activate Norms*, 34 PERSONALITY & SOC. PSYCHOL. BULL. 1047 (2008) (replicating and extending the Aarts & Dijksterhaus “silence in the library” study).

108. See Robert B. Cialdini et al., *A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places*, 58 J. PERSONALITY & SOC. PSYCHOL. 1015, 1016-17 (1990).

dane physical cues can also have a behavioral effect.¹⁰⁹ For example, primes such as a leather briefcase and a fancy pen have been shown to invoke the same behavioral effect as the lexical primes used to invoke the stereotype “businessperson.”¹¹⁰ Most recently, experimental subjects exhibited increased interpersonal warmth and a greater likelihood to behave generously when they were supraliminally primed by being asked to briefly hold a warm cup of coffee in the elevator on the way to the testing room.¹¹¹

Finally, research on the priming effect has shown that it is possible to affect an individual’s level of group identification by priming “I” versus “we” associations.¹¹² This is potentially very important, because the SIDE model of deindividuation demonstrates the importance of social group identification to either normative or antinormative behavior.¹¹³ The extent to which an individual is inclined to view a particular group as self-relevant can affect whether the individual’s attitude and behavior move toward, or away from, the perceived attitudes and behaviors of the group.¹¹⁴ Furthermore, it is possible to manipulate the individual’s degree of identification with the group.¹¹⁵

In one set of experiments, subjects were asked to write an essay about a day in the life of another student, alternately named “Tyrone Walker” or “Erik Walker” (a subtle prime for either African American or White).¹¹⁶ The students later took a test; those primed with the African-American name performed worse on the test, but the effect was much more pronounced for those who had been asked to write their essays in the first person than for those who wrote in the third person.¹¹⁷ A summary of the research in this area has noted:

[I]t is relatively easy to imagine naturalistic settings in which personal or social self-construals may be activated relatively spontaneously by environmental cues. Under circumstances in which one’s

109. See Aaron C. Kay et al., *Material Priming: The Influence of Mundane Physical Objects on Situational Construal and Competitive Behavioral Choice*, 95 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 83 (2004).

110. See *id.* at 90-91. In the experiment, the researcher alternately took the research questionnaires out of a leather briefcase and provided a business-style pen to the subject or took the questionnaires out of a backpack and provided a pencil. *Id.* at 90.

111. Lawrence E. Williams & John A. Bargh, *Experiencing Physical Warmth Promotes Interpersonal Warmth*, 322 SCIENCE 606, 606-07 (2008).

112. See, e.g., Diederik A. Stapel & Willem Koomen, *I, We, and the Effects of Others on Me: How Self-Construal Level Moderates Social Comparison Effects*, 80 J. PERSONALITY & SOC. PSYCHOL. 766 (2001).

113. See *supra* notes 74-77 and accompanying text.

114. See Dijksterhuis et al., *Effects of Priming*, *supra* note 17, at 86-95.

115. See *id.* at 93-95.

116. S. Christian Wheeler et al., *Think unto Others: The Self-Destructive Impact of Negative Racial Stereotypes*, 37 J. EXPERIMENTAL SOC. PSYCHOL. 173, 175-76 (2001).

117. See *id.* at 175-76; see also Dijksterhuis et al., *Seeing One Thing*, *supra* note 91, at 867-68 (demonstrating that subjects conformed behavior toward the general category of “professor” but away from the exemplar “Einstein”).

individual performance or responsibility is stressed, the personal self is activated, whereas situations that stress communal feelings, consensus, and togetherness activate the social self.¹¹⁸

Perhaps unsurprisingly, stress and competition are among the factors that can lead to a more individualistic mindset.¹¹⁹

IV. CONCLUSION: THE POWER OF THE SITUATION AND THE POWER OF THE GROUP

Considered together, these separate literatures on deindividuation and on priming suggest that it might be possible for institutions to promote prosocial behavior and to reduce hate speech by fostering identification with those social identity categories for which the normative structures would tend to discourage hate speech. Furthermore, the research suggests that this could be done in subtle—perhaps nearly invisible—ways. Paradoxically, a sense of anonymity—or deindividuation—is triggered by identification with a group. However, it is not the anonymity per se that leads to antinormative behaviors. Rather, the identification with a social group tends to foster attitudes and behaviors consonant with the norms of the particular group. Because attitudes, behaviors, and group identification can be primed by features in the social and physical environment, universities potentially could influence behavior (including speech) in prosocial directions by carefully attending to those features over which they have some degree of control, such as physical spaces, institutional culture, and social organization.

One possible objection to this proposal is that it evokes the specter of manipulation and thought control. Just as the prospect of subliminal advertising is objectionable because it seems to undermine free will and consumer choice, so too the prospect of subtle or unconscious priming may strike some as an illegitimate way to influence behavior. Perhaps the best response to this objection is that situations inevitably influence thoughts and behaviors; if the choice is between moving individuals in a prosocial or antisocial direction, it seems obvious that the former is preferable. Just as architects design public spaces so as to reduce criminal behavior and to promote communal spirit, universities should design institutional space—both social and physical—to promote tolerance, empathy, and cooperation. To take just one rather mundane example, much research has demonstrated the effect of mirrors in decreasing antinormative behavior.¹²⁰ Other studies suggest

118. Stapel & Koomen, *supra* note 112, at 778-79.

119. *See id.* at 779.

120. *See, e.g.,* S. Christian Wheeler et al., *Does Self-Consciousness Increase or Decrease Priming Effects? It Depends*, 44 J. EXPERIMENTAL. PSYCHOL. 882, 883 (2008) (summarizing and analyzing the prior research).

that inclusive (“we”) primes in the environment can serve to increase cooperative behavior and to decrease competitiveness. Indeed, a feature as simple as the color of a space can affect behavior.¹²¹ Though this Essay is merely exploratory and very preliminary, social psychology research on the influence of situational primes upon behavior and goals is a vibrant and fast-moving area that could no doubt inform institutional design on these questions.

Taking such steps would have the further salutary effect that when ambiguous incidents did occur, members of minority groups would be less likely to interpret them as racist or threatening.¹²² Furthermore, to the extent that negative stereotypes tend to depress motivation and achievement, the perception of a safe and welcoming environment will counteract this effect.¹²³

After studying the issue of campus hate speech in great depth, a special committee at the University of Pennsylvania recommended, among other things, that the model for freshman housing assignment be redesigned and that fraternity and sorority rush be delayed until after the first semester of the first year.¹²⁴ These recommendations grew out of recognition that students were self-segregating as they chose their residences and that the structure of student living could affect other aspects of university life.

These recommendations were met with great resistance on the part of many students, who argued that the identity-based houses provided crucial support for minorities and other marginalized social groups. This conflict is an aspect of a larger puzzle that is highlighted by the individuation and priming research: group identification can foster either prosocial or antisocial behavior, depending on the norms of the particular social group which is salient to the individual at a given moment. Furthermore, any individual is at once a member of myriad social groups; the process of group identification is fluid and dynamic.

121. See Andrew J. Elliot et al., *Color and Psychological Functioning: The Effect of Red on Performance Attainment*, 136 J. EXPERIMENTAL PSYCHOL. 154 (2007).

122. The water buffalo incident at the University of Pennsylvania was ambiguous. The African-American women who were the target of the student’s remark interpreted it as racist, but the speaker insisted (and the University later all but conceded) that in his native language the term was a generalized insult that had no racial connotation. See ALAN CHARLES KORS & HARVEY A. SILVERGLATE, *THE SHADOW UNIVERSITY: THE BETRAYAL OF LIBERTY ON AMERICA’S CAMPUSES* (1998).

123. See Valerie Purdie-Vaughns et al., *Social Identity Contingencies: How Diversity Cues Signal Threat or Safety for African Americans in Mainstream Institutions*, 94 J. PERSONALITY & SOC. PSYCHOL. 615, 628 (2008). Purdie-Vaughns and her colleagues found that where a potential employer was perceived as ethnically diverse, a “colorblind” value statement, which the researchers characterized as ambiguous in terms of whether it signaled a welcome atmosphere for minority employees, was not perceived as threatening. *Id.* In contrast, where brochure photos depicted a homogenous workplace, the same value statement was viewed as threatening. *Id.*

124. See FINAL REPORT, *supra* note 22.

A step in the direction of a solution, perhaps, would be to gather very specific data on hate speech incidents on college campuses in order to determine the situational features that tend to give rise to such behavior; for example, time of day, physical location, weekday versus weekend, or involvement of alcohol. Armed with such knowledge, universities could be more strategic in the way that they either encourage or discourage identification with various social identity groups in different settings. Though antisocial behaviors, including hate speech, are unlikely to disappear altogether, it seems likely (or at least possible) that institutions could do more to employ the insights derived from social and cognitive psychological research to nudge behavior in the desired direction.

