

TROJAN HORSES OF RACE

Jerry Kang

TABLE OF CONTENTS

INTRODUCTION.....	1491
I. RACIAL MECHANICS.....	1497
<i>A. Racial Schemas</i>	1498
1. Schemas Generally.....	1498
2. Racial Schemas.....	1499
3. Automaticity.....	1504
<i>B. Implicit Bias</i>	1506
1. The Problem: Opacity.....	1506
2. The Solution: Measuring Speed.....	1508
3. The Results: Pervasive Implicit Bias.....	1512
<i>C. Behavioral Consequences</i>	1514
1. Interpreting.....	1515
2. Performing.....	1519
3. Interacting.....	1523
4. Shooting.....	1525
<i>D. Objections</i>	1528
1. Correction Is Easy.....	1528
2. Correction Is Impossible.....	1531
<i>E. A Research Agenda</i>	1535
II. TROJAN HORSES.....	1539
<i>A. Tuning In to Broadcast</i>	1539
<i>B. Redefining the Public Interest</i>	1545
<i>C. Local News</i>	1549
1. Crime and Punishment.....	1549
2. Trojan Horse Viruses.....	1553
3. The Accuracy Objection.....	1563
<i>D. Virus Protection</i>	1567
1. Recoding the Public Interest.....	1568
2. Thought Experiments.....	1572
(a) Firewall: Capping Crime Stories.....	1572
(b) Disinfection: Public Service Announcements.....	1579
3. The Autonomy Objection.....	1584
CONCLUSION.....	1589

TROJAN HORSES OF RACE

*Jerry Kang**

Recent social cognition research has provided stunning evidence of implicit bias against various social categories. In particular, it reveals that most of us have implicit biases against racial minorities notwithstanding sincere self-reports to the contrary. These implicit biases have real-world consequence — in how we interpret actions, perform on exams, interact with others, and even shoot a gun. The first half of this Article imports this remarkable science into the law reviews and sets out a broad intellectual agenda to explore its implications. The second half examines where implicit bias comes from, and focuses on vicarious experiences with racial others mediated through electronic communications. This, in turn, raises a timely question of broadcast policy sparked by the FCC's controversial 2003 Media Ownership Order. There, the FCC repeatedly justified relaxing ownership rules by explaining how it would increase, of all things, local news. But local news is replete with violent crime stories prominently featuring racial minorities. Consumption of these images, the social cognition research suggests, exacerbates our implicit biases. In other words, as we consume local news, we download a sort of Trojan Horse virus that increases our implicit bias. Unwittingly, the FCC linked the "public interest" to racism. Potential responses, such as recoding the public interest and examining potential firewalls and disinfectants for these viruses, are discussed in light of psychological, political, and constitutional constraints.

* Visiting Professor of Law, Georgetown University Law Center; Professor of Law, UCLA School of Law. E-mail: kang@law.ucla.edu. Website: <http://jerrykang.net>. Special thanks to Mahzarin Banaji. For helpful comments and conversations, I thank Annette Almazan, Betsy Bartholet, Gary Blasi, Angela Campbell, Ron Chen, Julie Cohen, Buju Dasgupta, Chris Edley, Dick Fallon, Terry Fisher, Kim Forde-Mazrui, James Forman, Oscar Gandy, Steve Goldberg, Tony Greenwald, Lani Guinier, Mitu Gulati, Linda Hamilton Krieger, Emma Jordan, John Jost, Duncan Kennedy, Randy Kennedy, Sung Hui Kim, Russell Korobkin, Don Langevoort, Carrie Menkel Meadow, John Mikhail, Helen Nissenbaum, Mike Seidman, Margaret Shih, Joe Singer, Gerry Spann, Rip Verkerke, Fred Yen, and Kathy Zeiler.

Research assistance was provided by Angela Chan, Chester Day, and Michael Trinh, with smaller contributions from Sean Byrne, Justin Levinson, and Tom Moss. The Hugh and Hazel Darling Law Library at the UCLA provided world-class assistance, as always. Funding for this project was provided by UCLA Asian American Studies Center, UCLA Dean's Summer Research Fund, Harvard Law School, and Georgetown University Law Center.

This paper has been presented at the Berkman Center Cyberlaw Retreat, Cornell Law School, Georgetown Law Center, Harvard Law School, New York University's Department of Communication and Culture, and UCLA School of Law. I dedicate this piece to the teachers who have inspired me.

There is no immaculate perception.

— Commonly attributed to Nietzsche¹

You are what you eat.

— Nutritional maxim

In all fighting, the direct method may be used for joining battle, but indirect methods will be needed in order to secure victory.

— Sun Tzu²

INTRODUCTION

Consider the following studies, with an open mind.

Computer Crash. Social cognitionist John Bargh asked participants to count whether an even or odd number of circles appeared on a computer screen.³ After the 130th iteration, the computer was designed to crash, and the participants were told to start over. A hidden video camera recorded their reactions. Third-party observers then evaluated those recordings to measure participants' frustration and hostility. What neither participants nor observers knew was that for half the participants, a young Black male face was flashed subliminally⁴ before each counting iteration; for the other half, the face was White. As rated by the observers, those who had been shown the Black faces responded with greater hostility to the computer crash.⁵

Mugshot. Political scientists Frank Gilliam and Shanto Iyengar created variations of a local newscast: a control version with no crime story, a crime story with no mugshot, a crime story with a Black-suspect mugshot, and a crime story with a White-suspect mugshot.⁶

¹ Apparently, Nietzsche never actually said this, although he used the term "immaculate perception" (*unbefleckten Erkenntniss*) in *Thus Spoke Zarathustra* to disparage traditional views of knowledge. See FRIEDRICH NIETZSCHE, *THUS SPOKE ZARATHUSTRA*, reprinted in *THE PORTABLE NIETZSCHE* 100, 233–36 (Walter Kaufmann ed. & trans., 1954).

² SUN TZU, *THE ART OF WAR*, reprinted in *ROOTS OF STRATEGY: A COLLECTION OF MILITARY CLASSICS* 31 (Thomas R. Phillips ed., 1940).

³ 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, 238–39 (1996) (describing experimental procedure).

⁴ Sometimes, such primes are called "preconscious" instead of "subliminal." See, e.g., Susan T. Fiske, *Stereotyping, Prejudice, and Discrimination*, in 2 *THE HANDBOOK OF SOCIAL PSYCHOLOGY* 357, 365–66 (Daniel T. Gilbert et al. eds., 4th ed. 1998).

⁵ See Bargh et al., *supra* note 3, at 239. Hostility measures were scored on a scale of 0 to 10, with 10 being most hostile. The blind coders evaluated the White-primed participants at a hostility of $M = 2.13$; by contrast, they coded the Black-primed participants at $M = 2.79$. The difference was statistically significant at $p < 0.05$. *Id.* Readers who see this effect as small might consider that each face appeared on the screen for only about one-fortieth of a second. *Id.* at 238.

⁶ Franklin D. Gilliam, Jr. & Shanto Iyengar, *Prime Suspects: The Influence of Local Television News on the Viewing Public*, 44 *AM. J. POL. SCI.* 560, 563–67 (2000) (describing experimental procedure). Participants answered a preliminary questionnaire, which solicited basic demo-

The Black and White suspects were represented by the same morphed photograph, with the only difference being skin hue — thus controlling for facial expression and features.⁷ The suspect appeared for only five seconds in a ten-minute newscast; nonetheless, the suspect's race produced statistically significant differences in a criminal law survey completed after the viewing. Having seen the Black suspect, White participants showed 6% more support for punitive remedies than did the control group, which saw no crime story.⁸ When participants were instead exposed to the White suspect, their support for punitive remedies increased by only 1%, which was not statistically significant.⁹

Math Test. Social psychologist Margaret Shih asked Asian American women at Harvard University to take a hard math test.¹⁰ Before taking the exam, each participant answered a questionnaire designed to prime subtly different social identities: female (with questions relating, for example, to coed dormitory policy) or Asian (with questions relating, for example, to language spoken at home). A control group answered questions related to neutral topics, such as telecommunications usage.¹¹ As measured by an exit survey, these questions had no conscious impact on self-reports of test difficulty, self-confidence in math ability, the number of questions attempted, or how well participants thought they did.¹² Yet something happened implicitly. The group that had its Asian identity triggered performed best in accuracy (54%); the group that had no identity triggered came in second (49%); and the

graphic, political affiliation, and media habits data, prior to watching the newscasts and completed a detailed questionnaire gauging crime-related and racial attitudes after the newscast. *Id.* at 564. The crime-related attitudes that were measured were fear of crime, dispositional explanations for crime, and support for punitive criminal justice. *Id.* at 565. Racial attitudes were measured on both “old fashioned” and “new” racism scales. *See id.* at 566. The 2331 participants were residents of the Los Angeles metropolitan area. *Id.* at 564. Reflecting the demographics of the area, 53% of the participants were White, 22% Black, 10% Asian, and 8% Latino. Fifty-two percent were women, 49% had graduated from college, 45% were Democrats, and 25% were Republicans. *Id.*

⁷ *Id.* at 563. The researchers separately validated their digitally produced pictures as mapping to the racial categories White and Black. *See id.* at 563 n.5.

⁸ *Id.* at 567–68 & tbl.3. Following the newscast, the researchers asked three questions about participants' support for certain criminal remedies: death penalty for murderers, three strikes legislation, and increased police presence on the streets. *Id.* at 565. Participants' responses were then scaled to an index between 0 and 1, with 1 being most punitive, to generate a punitive index. *See id.* at 568 tbl.3. The mean punitive index of the Black-suspect condition group was 0.06 higher than that of the control group. This result was statistically significant at $p < 0.01$. *Id.* at 568 & tbl.3.

⁹ *See id.* at 568 tbl.3.

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

¹¹ *Id.*

¹² *Id.* at 81.

group that had its female identity triggered ranked last (43%).¹³ “Being” *Asian* boosted, while “being” *female* depressed, math performance. Of course, these students were both.

Shooter Bias. Social cognitionist Joshua Correll created a video game that placed photographs of a White or Black individual holding either a gun or other object (wallet, soda can, or cell phone) into diverse photographic backgrounds.¹⁴ Participants were instructed to decide as quickly as possible whether to shoot the target. Severe time pressure designed into the game forced errors. Consistent with earlier findings,¹⁵ participants were more likely to mistake a Black target as armed when he in fact was unarmed (false alarms);¹⁶ conversely, they were more likely to mistake a White target as unarmed when he in fact was armed (misses).¹⁷ Even more striking is that Black participants showed similar amounts of “shooter bias” as Whites.¹⁸

What is going on here? Quite simply, a revolution. These studies are the tip of the iceberg of recent social cognition research elaborating what I call “racial mechanics”¹⁹ — the ways in which race alters intrapersonal, interpersonal, and intergroup interactions. The results are stunning, reproducible, and valid by traditional scientific metrics. They seriously challenge current understandings of our “rational” selves and our interrelations.

In Part I, I import crucial findings from the field of social cognition with emphasis on the recent “implicit bias” literature.²⁰ This research

¹³ *Id.* The study defined “accuracy” as the number of questions a participant answered correctly divided by the number of questions attempted. *See id.* A linear contrast analysis on this pattern of scoring from top to bottom (Asian identity group to control group to female identity group) was statistically significant at $p < 0.05$. *Id.* When the measure was changed from accuracy to absolute number of questions answered correctly, the same pattern appeared but statistical significance was lost. *Id.* Out of twelve questions, the Asian identity group answered $M = 5.37$ correctly; the control group $M = 5.31$; the female identity group $M = 4.71$ ($p = 0.19$). *Id.*

¹⁴ Joshua Correll et al., *The Police Officer's Dilemma: Using Ethnicity To Disambiguate Potentially Threatening Individuals*, 83 J. PERSONALITY & SOC. PSYCHOL. 1314, 1315–17 (2002) (describing experimental setup).

¹⁵ *See* B. Keith Payne, *Prejudice and Perception: The Role of Automatic and Controlled Processes in Misperceiving a Weapon*, 81 J. PERSONALITY & SOC. PSYCHOL. 181, 185–86 (2001).

¹⁶ Correll et al., *supra* note 14, at 1319. This finding was statistically significant at $p < 0.02$; however, when outlier target images were put back into the dataset, the significance disappeared. The researchers removed certain outlier images because they produced so many errors, suggesting that something in the background image or person’s clothing produced misleading impressions. *Id.*

¹⁷ *Id.* This finding was significant at $p < 0.001$. *Id.*

¹⁸ *Id.* at 1325.

¹⁹ *See* Jerry Kang, *Cyber-Race*, 113 HARV. L. REV. 1130, 1138–46 (2000) (introducing the “racial mechanics” model).

²⁰ Ron Chen and Jon Hanson have recently authored a fine and comprehensive review of the social cognition literature, although not with special emphasis on “implicit bias.” *See* Ronald Chen & Jon Hanson, *Categorically Biased: The Influence of Knowledge Structures on Law and Legal Theory*, 77 S. CAL. L. REV. 1103 (2004).

demonstrates that most of us have implicit biases in the form of negative beliefs (stereotypes) and attitudes (prejudice) against racial minorities. These implicit biases, however, are not well reflected in explicit self-reported measures. This dissociation arises not solely because we try to sound more politically correct. Even when we are honest, we simply lack introspective insight. Finally, and most importantly, these implicit biases have real-world consequences — not only in the extraordinary case of shooting a gun, but also in the more mundane, everyday realm of social interactions.

A vast intellectual agenda opens when we start probing what this new knowledge might mean for law. Part I, which is necessarily long and detailed, establishes the foundation for that scholarship in the law reviews.²¹ In Part II, I focus on a single application of the racial mechanics model to a timely question of communications policy. I start by asking a fundamental question: “Where does bias come from?” One important source is vicarious experience with the racial other, transmitted through the media. If these experiences are somehow skewed, we should not be surprised by the presence of pervasive implicit bias. What, then, might we do about such media programming given the rigid constraints of the First Amendment? To be sure, private actors of good faith can voluntarily adopt best practices that decrease implicit bias and its manifestations. But can the state, through law, do anything?

If there is any room for intervention, it would be in the communications realm of broadcast, which enjoys doctrinal exceptionalism. In broadcast, notwithstanding the First Amendment, we tolerate the licensing of speakers. In broadcast, we tolerate suppression of speech we dislike, such as indecency and violence. In broadcast, we tolerate

²¹ This research is different from, but broadly consistent with, the cognitive psychological groundings of behavioral decision theory, or “behavioral law and economics.” Don Langevoort provides an excellent literature review. See Donald C. Langevoort, *Behavioral Theories of Judgment and Decision Making in Legal Scholarship: A Literature Review*, 51 VAND. L. REV. 1499 (1998); see also Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471 (1998).

The implicit bias social cognition research I describe is both narrower and broader than behavioral decision theory as practiced in the law reviews. Its focus is narrower because instead of all sorts of decisions, it focuses on those influenced by or made across social categories such as race and gender. At the same time, the implicit bias research feels broader in its ambitions. The project is not merely to identify limited perturbations from a starting point of Rational Choice Theory (RCT). Rather, as Part I relates, the implicit bias research suggests far more pervasive, fundamental, and arational cognitive processes such that RCT, even with modifications, fails to model adequately the functioning of bias. Indeed, it is time for a new “behavioral realist” approach, which draws on the traditions of legal realism and behavioral science. This term was recently coined in an informal discussion among Mahzarin Banaji, Gary Blasi, Anthony Greenwald, John Jost, Linda Hamilton Krieger, and me to name an ongoing collaboration of legal academics and social cognitionists that seeks to apply the best model of human behavior that science has made available to questions of law and policy.

encouragement of speech we like, such as educational television and local-oriented programming. All this is in the name of the “public interest,” the vague standard that Congress has charged the Federal Communications Commission with pursuing.

That “public interest” standard was recently reshaped in the controversial June 2003 Media Ownership Order.²² There, the FCC repeatedly justified relaxing ownership rules by explaining how such changes would increase, of all things, local news.²³ Since local news was viewed as advancing “diversity” and “localism,” two of the three core elements of the “public interest,”²⁴ any structural deregulation that would increase local news was lauded.

Troubling is what’s on the local news. Sensationalistic crime stories are disproportionately shown: “If it bleeds, it leads.” Racial minorities are repeatedly featured as violent criminals. Consumption of these images, the social cognition research suggests, exacerbates our implicit biases against racial minorities. Here, then, is the link between Parts I and II: since implicit bias is fueled in part by what we see, the FCC has recently redefined the public interest so as to encourage the production of programming that makes us more biased. Unwittingly, the FCC linked the public interest to racism. No one spotted the issue for the Commission.

For a race paper, my using social cognition and applying it to communications law are unorthodox, but purposefully so. Race talk in legal literature feels like it is at a dead end. No new philosophical argument or constitutional theory seems to persuade those sitting on one side of the fence to jump to the other. One way to break current deadlocks is to turn to new bodies of knowledge uncovered by social science,²⁵ specifically the remarkable findings of social cognition.²⁶ Not only do they provide a more precise, particularized, and empirically grounded picture of how race functions in our minds, and thus in our societies, they also rattle us out of a complacency enjoyed after the demise of de jure discrimination.²⁷ Further calls for equality are often

²² In the Matter of 2002 Biennial Regulatory Review, 18 F.C.C.R. 13,620 (2003) [hereinafter Media Ownership Order], *aff’d in part and remanded*, Prometheus Radio Project v. FCC, 373 F.3d 372 (3d Cir. 2004).

²³ See *id.* at 13,643–45. For references to local news, see *infra* section II.B, pp. 1545–49.

²⁴ The third element is “competition.” Media Ownership Order, *supra* note 22, at 13,627.

²⁵ I consider the remarkable empirical, econometric, and game-theoretic work of Ian Ayres to be a sterling example of this turn. See generally IAN AYRES, PERVASIVE PREJUDICE? UNCONVENTIONAL EVIDENCE OF RACE AND GENDER DISCRIMINATION 12–13 (2001) (describing how econometrics and game theory are underutilized tools for analyzing civil rights).

²⁶ An accessible introductory volume to social cognition is SUSAN T. FISKE & SHELLEY E. TAYLOR, SOCIAL COGNITION (2d ed. 1991).

²⁷ Cf. Ian Ayres, *Is Discrimination Elusive?*, 55 STAN. L. REV. 2419, 2428 (2003) (book review) (identifying the importance of empirical evidence that cannot be as easily dismissed by “the general public and lawmakers” as personal narrative can).

derogated as whining by those who cannot compete in a modern meritocracy. Social cognition discoveries dispute that resentful characterization and make us reexamine our individual and collective responsibilities for persistent racial inequality. The inquiry is not, however, a simple witch hunt by those “pure” for those “tainted”: as Charles Lawrence observed long ago, albeit drawing on a very different psychology, we may all be infected in ways we cannot admit, even to ourselves.²⁸

My point is not that “hard science” is the only way to engage in critical thinking about the relationship between law and race.²⁹ This would respark an unhelpful methodology war we endured in the 1980s and 1990s about the essential role of narrative in critical race theory.³⁰ Like most scholars, I believe that multiple methodologies produce the deepest insight.³¹ I am also aware of the limits of the scientific method and of the ignominious history of pseudoscience’s complicity in brutally subordinating entire peoples.³² That said, I am confident that the language, methodologies, and findings of social cognition provide trenchant additions to the philosophical, anthropological, sociological,

²⁸ This recognition and reframing could help us get beyond “blaming” and toward “solving.” Cf. Charles R. Lawrence III, *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317, 321 (1987) (suggesting the metaphor of illness for the problem of racism). In this seminal article, Charles Lawrence argues that the “cultural meaning” of allegedly racially discriminatory action is the best evidence of an unconscious collective racial bias. *Id.* at 323–24. Lawrence’s work is consistent with the claims of this Article, broadly stated. Moreover, its exploration of unconscious racism, based on psychology, is a crucial intellectual precursor. However, social cognition is a very different psychology from that of Freud and psychoanalysis. See *id.* at 332 (drawing on psychoanalytic theory). Also, Lawrence’s methodology divines cultural meaning, whereas the “implicit bias” research measures bias through individuals’ reaction-time differentials.

²⁹ If more evidence of sincerity is necessary, see, for example, Jerry Kang & Benedikt Buchner, *Privacy in Atlantis*, 18 HARV. J.L. & TECH. 229 (2004), which is written entirely as a Socratic dialogue, with five fictional characters. See also Kang, *supra* note 19, at 1135, 1140 (employing both nonfictional and fictional stories).

³⁰ See, e.g., Daniel A. Farber & Suzanna Sherry, *Telling Stories Out of School: An Essay on Legal Narratives*, 45 STAN. L. REV. 807 (1993) (criticizing various forms of narrative scholarship); Mark Tushnet, *The Degradation of Constitutional Discourse*, 81 GEO. L.J. 251 (1992) (same). In the field of Asian American Jurisprudence, see the extraordinary conflict between Robert S. Chang, *Toward an Asian American Legal Scholarship: Critical Race Theory, Post-Structuralism, and Narrative Space*, 81 CAL. L. REV. 1241 (1993), and Jim Chen, *Unloving*, 80 IOWA L. REV. 145 (1994). The resulting colloquy appears in the *Iowa Law Review*. See, e.g., Keith Aoki, *The Scholarship of Reconstruction and the Politics of Backlash*, 81 IOWA L. REV. 1467 (1996); Jim Chen, *Response: Untenured but Unrepentant*, 81 IOWA L. REV. 1609 (1996); Margaret Chon, *Chon on Chen on Chang*, 81 IOWA L. REV. 1535 (1996); Neil Gotanda, *Chen the Chosen: Reflections on Unloving*, 81 IOWA L. REV. 1585 (1996); Alfred C. Yen, *Unhelpful*, 81 IOWA L. REV. 1573 (1996). For a thoughtful and more measured critique of Chang’s call for postmodernism, see Gary Chartier, *Righting Narrative: Robert Chang, Poststructuralism, and the Possibility of Critique*, 7 UCLA ASIAN PAC. AM. L.J. 105 (2001).

³¹ Cf. Ayres, *supra* note 27, at 2427 (calling himself an “empirical pluralist”); Devon W. Carbado & Mitu Gulati, *The Law and Economics of Critical Race Theory*, 112 YALE L.J. 1757, 1785 (2003) (book review) (noting that narrative does not preclude statistical analysis).

³² See generally STEPHEN JAY GOULD, *THE MISMEASURE OF MAN* (1981).

literary, and political science modes of argument that have so far dominated critical race studies.³³ For better and worse, law has turned sharply in favor of quantified and empirical analyses. Social cognition allows a phalanx of those who study race to take that same turn, instrumentally to fight fire with fire, and substantively to profit from a body of science that supports, particularizes, and checks what we intuit as the truth of our lived experiences.³⁴ The potential reward in insight and persuasion is substantial, as I hope to demonstrate.

Another way to generate new insights is to view old topics through new lenses.³⁵ That explains my invocation of the metaphor of “Trojan Horses,” which is more familiar to cyberlaw than to critical race studies. This strategy further explains why I apply my social cognitive model of racial mechanics to FCC regulations.³⁶ I start with the theory and evidence of racial mechanics.

I. RACIAL MECHANICS

I have coined the term “racial mechanics” to describe how race alters interpersonal interactions.³⁷ My model draws heavily from the field of social cognition, with emphasis on the recent implicit bias literature. For most lawyers and legal academics, this science will be jaw-dropping.³⁸ For social cognitionists, what will be eye-opening is the theoretical translation of social cognitive findings to themes in

³³ For an introduction to Critical Race Theory, see *CRITICAL RACE THEORY: THE KEY WRITINGS THAT FORMED THE MOVEMENT*, at xiii–xxxii (Kimberlé Crenshaw et al. eds., 1995). See also Richard Delgado & Jean Stefancic, *Critical Race Theory: An Annotated Bibliography 1993, A Year of Transition*, 66 U. COLO. L. REV. 159 (1994) (outlining eleven themes of critical race theory).

³⁴ Various legal scholars studying race have already made helpful empirical and quantitative turns. Consider, for instance, the trailblazing work of Ian Ayres. See AYRES, *supra* note 25; see also Pat K. Chew, *Asian Americans in the Legal Academy: An Empirical and Narrative Profile*, 3 ASIAN L.J. 7 (1996); William C. Kidder, *Situating Asian Pacific Americans in the Law School Affirmative Action Debate: Empirical Facts About Thernstrom's Rhetorical Acts*, 7 ASIAN L.J. 29 (2000); Alfred C. Yen, *A Statistical Analysis of Asian Americans and the Affirmative Action Hiring of Law School Faculty*, 3 ASIAN L.J. 39 (1996).

³⁵ See, e.g., Carbado & Gulati, *supra* note 31 (applying organizational management and economics insights to critical race theory).

³⁶ This choice in topic should not be read as an implicit claim that “local news” is the most obvious, compelling, or convincing application of the social science I could find. To the contrary, I selected a nonobvious, particularly challenging application in order to probe the theory’s potential and to advance my broader project of cross-fertilizing race and communications scholarship. See, e.g., Kang, *supra* note 19 (discussing social construction of race within the technological construction of cyberspace); Jerry Kang, *E-racing E-lections*, 34 LOY. L.A. L. REV. 1155 (2001) (discussing race and electronic voting).

³⁷ See Kang, *supra* note 19, at 1138–47.

³⁸ Academics in the behavioral law and economics tradition may be less surprised, but even they should be impressed by the reaction-time measurements of implicit bias. For a tentative mapping of the relationship between that literature and the social science described here, see *supra* note 21.

critical race studies and the practical translation to potential legal and policy reforms.

A. Racial Schemas

1. *Schemas Generally.* — A schema is a “cognitive structure that represents knowledge about a concept or type of stimulus, including its attributes and the relations among those attributes.”³⁹ These knowledge structures can be modeled as prototypes or exemplars for a class of objects, providing rules that map objects into the class, as well as general information about members of the class.⁴⁰ For instance, when we see something that has four legs, a horizontal plane, and a back, we immediately classify that object into the category “chair.” We then un-

³⁹ FISKE & TAYLOR, *supra* note 26, at 98. There are competing definitions for similar ideas in the literature. For example, Ziva Kunda proposes the term “concept” instead, which she defines as “a mental representation of a category, that is, a class of objects that we believe belong together.” ZIVA KUNDA, *SOCIAL COGNITION: MAKING SENSE OF PEOPLE* 16 (1999). She notes that “concepts” are often called “schemas” by social psychologists. *See id.* at 17; *see also* MARTHA AUGOUSTINOS & IAIN WALKER, *SOCIAL COGNITION: AN INTEGRATED INTRODUCTION* 26 (1995) (“A schema is a mental structure of some referent which consists of knowledge and examples of that referent, and which selects and processes information pertinent to that referent.”); Charles Stangor & Mark Schaller, *Stereotypes as Individual and Collective Representations*, in *STEREOTYPES AND STEREOTYPING* 3, 7 (C. Neil Macrae et al. eds., 1996) (“Schemas are abstract knowledge structures that specify the defining features and relevant attributes of a given concept.”).

⁴⁰ My summary submerges controversy about the best ways to understand schemas. The classical view understood schemas as sharply “defined by a set of necessary and sufficient attributes.” *See* KUNDA, *supra* note 39, at 25. This formal, overly rigid model has fallen into disrepute. *See id.* In its place have come two alternative models, the prototype model and the exemplar model, both of which emphasize the probabilistic nature of schemas.

Instead of having necessary and sufficient definitional attributes, the prototype model emphasizes a core conception of the category that is created by some summary or average representation of every element in that category. *See id.* at 30. Whether some object belongs to that category is answered in terms of “family resemblance,” with a smooth variation from the objects that perfectly match the prototype to objects that seem more different.

By contrast, the exemplar model rejects the idea that the meanings associated with a category are somehow blended together into a single prototype. Instead, it contends that the racial meanings are connected to multiple examples or instances of the category that have been encountered in the past. *See id.*

Finally, a more recent theory-based conception of the schema suggests that a schema is to an instance as a theory is to data. The basic justification for such an approach is that probabilistic definitions cannot explain why certain similarities are considered relevant for categorization, while others are not. *See id.* at 37, 99.

Even this elaboration supposes a greater consensus on definitions than exists in the literature. There is pervasive equivocation of terms; some scholars, for example, reserve the word “schema” for the classic conception only and use the word “prototype” for what I am calling the prototype model of a schema. Lawyers entering this debate will question the value of these various definitional disputes as they smack of formalism. *See, e.g.,* Stangor & Schaller, *supra* note 39, at 16 (pointing out that the classical, prototype, and exemplar models are hard to distinguish).

derstand how to use the object, for example, by sitting on it. This schematic thinking operates automatically, nearly instantaneously.⁴¹

We employ schemas out of necessity. Our senses are constantly bombarded by environmental stimuli, which must be processed, then encoded into memories (short- and/or long-term) in some internal representation. Based on that representation of reality, we must respond. But we drown in information. Perforce⁴² we simplify the datastream at every stage of information processing through the use of schemas.⁴³

Different schema types exist for different types of entities, such as objects, other people, the self, roles, and events.⁴⁴ To be clear, this most basic process operates not only on inanimate objects, such as chairs or bananas,⁴⁵ but also on human beings.⁴⁶ When we encounter a person, we classify that person into numerous social categories, such as gender, (dis)ability, age, race, and role.⁴⁷ My focus is on race.

2. *Racial Schemas.* — Through law and culture, society provides us (the perceivers) with a set of racial *categories* into which we map an individual human being (the target) according to prevailing rules of racial *mapping*. Once a person is assigned to a racial category, implicit and explicit racial meanings associated with that category are triggered. These activated racial *meanings* then influence our interpersonal interaction. All three elements (presented as ovals in Figure 1)

⁴¹ Although “schemas” are fundamental to my analysis, they may be best understood as metaphor. More recent connectionist models of mental processes may explain empirical results as well as or better than schema-based approaches and also may be more parsimonious. See, e.g., Anthony G. Greenwald et al., *A Unified Theory of Implicit Attitudes, Stereotypes, Self-Esteem, and Self-Concept*, 109 PSYCHOL. REV. 3 (2002) (providing a unified theory of fundamental concepts in social cognition through a connectionist model similar to Heider’s balance theory). In Greenwald’s model, the basic analytical primitives are the “concept” (which includes persons, groups, and attributes), “association strength” (which measures the potential for one concept to activate another), and “concept activation” (which can take place by external stimuli or excitation through association with already activated concepts). See *id.* at 4–5. In future work, I hope to import these connectionist models into the legal literature.

⁴² See, e.g., Fiske, *supra* note 4, at 367 (describing categorical thinking as necessary to preserve mental resources and provide a sort of “cognitive economy”).

⁴³ According to Susan Fiske and Shelley Taylor, in the field of social cognition, there have historically been four different general views of the thinker: the consistency seeker (avoider of cognitive dissonance), naïve scientist (quasi-scientific analyzer of causes and effects), cognitive miser (seeker of rapid, adequate solutions or responses to situations), and motivated tactician (selective employer of various cognitive strategies, as driven by goals and motives). See FISKE & TAYLOR, *supra* note 26, at 10–13. Here I am emphasizing the view of the thinker in the last two senses.

⁴⁴ See *id.* at 118–20.

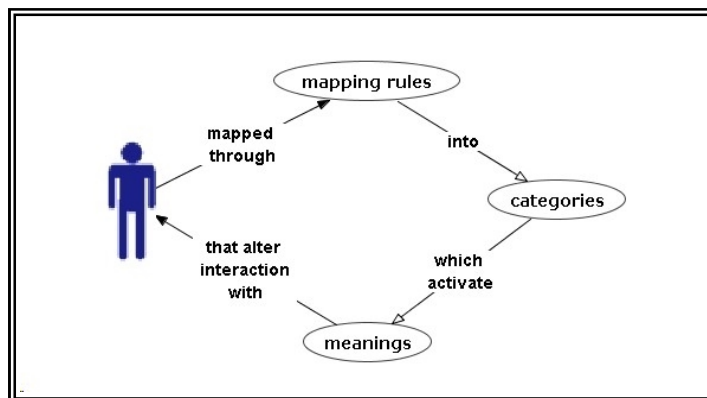
⁴⁵ See KUNDA, *supra* note 39, at 18 (providing banana example).

⁴⁶ See FISKE & TAYLOR, *supra* note 26, at 18–19 (noting differences between interpreting people and objects, which include the facts that people perceive back, can alter their behavior because they are being observed, and are often more complicated than objects).

⁴⁷ There is evidence that we usually acquire social categories in this order: gender, age, then race. See Fiske, *supra* note 4, at 376.

— racial categories, racial mapping rules, and racial meanings — constitute components of the racial schema.

FIGURE 1



I include in the term “racial meanings” both cognitive and affective components. The *cognitive* component includes thoughts or beliefs about the category, such as generalizations about their intelligence or criminality. The *affective* component reflects emotions, feelings, and evaluations that range on the scales of positive/negative, good/bad, approach/avoid.⁴⁸ Social psychologists often call cognitive beliefs about social groups “stereotypes”⁴⁹ and (negative) affective feelings about such groups “prejudice.”⁵⁰ The racial meanings triggered upon schema

⁴⁸ In the psychological literature, the affective component is often called an “attitude.” Attitudes generally emphasize evaluation. See, e.g., FISKE & TAYLOR, *supra* note 26, at 492 (giving one definition of attitude as an “association in memory between a given object and one’s evaluation of that object”). Some social psychologists use “attitude” as almost an alternative to my use of “schema.” For example, Mark Zanna and John Rempel define “attitude” as “the categorization of a stimulus object along an evaluative dimension based upon, or generated from, three general classes of information: (1) cognitive information, (2) affective/emotional information, and/or (3) information concerning past behaviors or behavioral intentions.” Mark P. Zanna & John K. Rempel, *Attitudes: A New Look at an Old Concept*, in THE SOCIAL PSYCHOLOGY OF KNOWLEDGE 315, 319 (Daniel Bar-Tal & Arie W. Kruglanski eds., 1988). Note again the emphasis on evaluation. See also AUGOUSTINOS & WALKER, *supra* note 39, at 12 (noting lack of clear consensus definition for “attitude”).

⁴⁹ Stereotypes have been defined in numerous ways in the literature. Based on an early definition by Walter Lippmann, many refer to stereotypes as “pictures in the head.” See, e.g., Stangor & Schaller, *supra* note 39, at 3 (quoting WALTER LIPPMAN, PUBLIC OPINION 3 (1922)).

⁵⁰ See, e.g., Fiske, *supra* note 4, at 357. This vocabulary is not universally employed. See, e.g., RUPERT BROWN, PREJUDICE: ITS SOCIAL PSYCHOLOGY 8 (1995) (defining prejudice as “the holding of derogatory social attitudes or cognitive beliefs, the expression of negative affect, or the display of hostile or discriminatory behaviour towards members of a group on account of their membership of that group”).

activation include both thoughts and feelings, rational and emotional, that can independently and jointly drive the perceiver's reactions.⁵¹

This social cognitive framing of racial mechanics provides new, concrete translations of various themes central to critical race studies. For example, critical race scholars repeat the mantra that "race is a social construction."⁵² My social cognitive account provides a particularized understanding of that general claim: all three components — racial categories,⁵³ mapping rules,⁵⁴ and racial meanings⁵⁵ — are

Legal commentators have not always used these terms of art in the same way the social scientists do. For example, Jody Armour characterizes "stereotypes" as automatic internal associations (what I would call "implicit" processes) and "prejudice" as self-conscious, chosen beliefs (what I would call "controlled" processes). See Jody Armour, *Stereotypes and Prejudice: Helping Legal Decisionmakers Break the Prejudice Habit*, 83 CAL. L. REV. 733, 749–50 (1995).

⁵¹ See, e.g., Fiske, *supra* note 4, at 357. There may be differences between affective and cognitive responses. For instance, self-reports of bias as well as instrumental action may be driven by cognitive processes, whereas noninstrumental behavior, especially if forced to be made quickly in response to environmental stimuli, may be driven by affective processes. See, e.g., FISKE & TAYLOR, *supra* note 26, at 522.

⁵² See, e.g., Cheryl I. Harris, *Equal Treatment and the Reproduction of Inequality*, 69 FORDHAM L. REV. 1753, 1759–64 (2001); Ian F. Haney López, *The Social Construction of Race: Some Observations on Illusion, Fabrication, and Choice*, 29 HARV. C.R.-C.L. L. REV. 1, 10–39 (1994). Much of the legal analysis in this vein draws from the sociological work of Michael Omi and Howard Winant. See MICHAEL OMI & HOWARD WINANT, *RACIAL FORMATION IN THE UNITED STATES: FROM THE 1960S TO THE 1990S* (2d ed. 1994).

⁵³ The variety and contours of racial categories have changed over time in American culture and law. See John A. Powell, *A Minority-Majority Nation: Racing the Population in the Twenty-First Century*, 29 FORDHAM URB. L.J. 1395, 1401–02 (2002) (suggesting that census categories have been altered in ways that reflect and maintain White supremacy).

⁵⁴ Rules of racial mapping have also been in flux. The law has had famously difficult times mapping multiracial individuals. An example of mapping rules in flux at the group level is what is happening in a post-9/11 world to the emerging, inchoate category of Middle Easterners/Arabs/Muslims, many of whom had previously fallen into the racial category of White. See Karen Engle, *Constructing Good Aliens and Good Citizens: Legitimizing the War on Terror(ism)*, 75 U. COLO. L. REV. 59, 72–77 (2004); see also John Tehranian, *Performing Whiteness: Naturalization Litigation and the Construction of Racial Identity in America*, 109 YALE L.J. 817, 837–41, 843–45 (2000) (documenting the history of conflicted judicial definitions of Arabs as both White and non-White under immigration and discrimination laws).

⁵⁵ By the end of the nineteenth century in America, for example, Asians (mostly Chinese) were seen as lecherous, vermin-like illegal immigrants who debased White society and labor — a far cry from the more familiar post-1960s "model minority." See, e.g., *Chae Chan Ping v. United States*, 130 U.S. 581, 595 (1889) ("As [Chinese immigrants] grew in numbers each year the people of the coast saw, or believed they saw . . . great danger that at no distant day that portion of our country would be overrun by them unless prompt action was taken to restrict their immigration."); *People v. Hall*, 4 Cal. 399, 405 (1854) (describing the Chinese as a people "whose mendacity is proverbial; a race of people whom nature has marked as inferior, and who are incapable of progress or intellectual development beyond a certain point, as their history has shown; differing in language, opinions, color, and physical conformation; between whom and ourselves nature has placed an impassable difference"). As I have previously noted, however, the "model minority" is tinged with the trope of "unfair competitor," which in turn connects back to historical attitudes toward the nineteenth-century Chinese laborer. See Note, *Racial Violence Against Asian Americans*, 106 HARV. L. REV. 1926, 1932 n.43 (1993). See generally Keith Aoki, "Foreign-ness" and

contingent, constructed, and contestable. Not one of these elements is biologically inevitable.

This framing also helps translate the critical race theme “race (almost always) matters.” In interpersonal encounters, multiple schemas may be activated.⁵⁶ For example, when one first meets a young Asian male law professor, multiple schemas could come into play, including age, race, gender, and role (profession). Which schemas actively influence the interaction depend on numerous variables, such as primacy (what gets activated first),⁵⁷ salience (which schema cues catch attention),⁵⁸ accessibility (which schemas can be retrieved in memory easily, perhaps because of recent priming),⁵⁹ and individuating information.⁶⁰

Notwithstanding such complexity and the variance among perceivers and environments,⁶¹ the scientific consensus is that racial schemas

Asian American Identities: Yellowface, World War II Propaganda, and Bifurcated Racial Stereotypes, 4 ASIAN PAC. AM. L.J. 1 (1996) (describing the history of Asian-American stereotypes).

⁵⁶ In order for a schema to influence information processing, it must somehow be “activated.” See, e.g., KUNDA, *supra* note 39, at 22 (describing activation as bringing some concept to mind).

⁵⁷ See, e.g., FISKE & TAYLOR, *supra* note 26, at 145 (discussing “primacy” and how the initial information we acquire influences the organization of subsequent information). Visual information tends to disclose age, gender, and race immediately and simultaneously. But one could have been told before meeting a person that he was a law professor, in which case the role schema would have been activated first.

⁵⁸ See *id.* In a room full of law professors, role may not be as salient as race. At a social gathering of Asian Americans, on the other hand, race may not be as salient as role.

⁵⁹ See *id.* Priming “refers to the incidental activation of knowledge structures, such as trait concepts and stereotypes, by the current situational context.” Bargh et al., *supra* note 3, at 230. Also, there is the possibility of schemas canceling each other out on some relevant metric, for example if the target is simultaneously a member of one ingroup and one outgroup. See, e.g., BROWN, *supra* note 50, at 49–54 (discussing cross-categorization and providing some evidence of canceling out as well as examples in which particular social categories predominate, such as religion over nationality among Muslims/Hindus and Indians/Bangladeshis).

⁶⁰ Other variables include mood, motivation (for example, for accuracy), hierarchy (for example, we tend to pay more careful, individuating attention to those who wield power over us), and cognitive busyness (or “load”). See FISKE & TAYLOR, *supra* note 26, at 145–467. Cognitive load can prevent the initial activation of a schema, but once the schema is activated, the load can increase reliance on the schema since the schema conserves mental resources. See JAMES M. JONES, PREJUDICE AND RACISM 201 (2d ed. 1997).

⁶¹ See, e.g., Russell H. Fazio & Bridget C. Dunton, *Categorization by Race: The Impact of Automatic and Controlled Components of Racial Prejudice*, 33 J. EXPERIMENTAL SOC. PSYCHOL. 451, 452 (1997) (classifying variables in terms of the environment and the perceiver). In this study, the researchers first measured participants for their implicit bias (automatic process) as well as their motivation to control prejudiced reactions (controlled process). See *id.* at 456–58. Then, participants were required to decide quickly whether two pictures were “similar” to each other; as between the pictures, race, gender, or profession was altered. See *id.* Those participants with intense automatic racial meanings, either positive or negative, used race more heavily in their similarity judgments. See *id.* at 461–62. Those who were more highly motivated to check their prejudice shifted attention away from race and toward occupation. See *id.* at 462. As the trials went on, almost every group — high implicit bias/low motivation to check, low implicit bias/low motivation to check, low implicit bias/high motivation to check — got faster in completing the similarity sorting tasks. See *id.* at 466. The one group whose speed did not increase was composed of those who had high implicit bias but also high motivation to check. See *id.*

are not of minor significance. Instead, racial schemas are “chronically accessible” and can be triggered by the target’s mere appearance, since we as observers are especially sensitive to visual and physical cues.⁶² As the Computer Crash study demonstrated, the visual information can be so fleeting as to be subliminal, yet it can still activate the racial schema.⁶³ We may not be colorblind even when we cannot see.

Once activated, the racial meanings embedded within the racial schema influence interaction. The apocryphal quotation attributed to Nietzsche, that “there is no immaculate *perception*,”⁶⁴ nicely captures how schemas guide what we see, encode into memory, and subsequently recall. At the attentional stage, schemas influence what we notice and immediately reduce information complexity.⁶⁵ At the encoding⁶⁶ and recall phases, schemas are again influential, although the memory literature is conflicted and qualified.⁶⁷ There is now evidence that schemas influence not only interpretation (that is, “social percep-

⁶² See FISKE & TAYLOR, *supra* note 26, at 144 (noting that “people rely heavily on visual, physical cues”); *id.* at 121 (noting that “certain physical characteristics may act almost like schematic labels”); *id.* at 145 (noting the importance of “salience,” in which people use schemas that catch attention); see also KUNDA, *supra* note 39, at 22 (suggesting that both looks and name will trigger racial schemas); Fiske, *supra* note 4, at 375 (noting that categories communicated visually have more impact than those communicated verbally).

⁶³ This is not to say that race is always or necessarily the first or most important schema to be activated. There is good evidence that people use “role” schemas first. See, e.g., FISKE & TAYLOR, *supra* note 26, at 143. Accordingly, when we encounter a police officer, for example, the role of “law enforcer” may be primary to gender or race, even if gender and race play an important role in the interaction.

⁶⁴ See *supra* note 1. Gestalt psychologists recognized early on that our perceptions are hardly unfiltered. They pointed out that perception is both constructed and mediated by the interpretive faculties of our central nervous system. See FISKE & TAYLOR, *supra* note 26, at 99.

⁶⁵ See Stangor & Schaller, *supra* note 39, at 6–7.

⁶⁶ If schemas are reasonably well-formed, people take longer to encode inconsistent information. By contrast, if schemas are weak, people may be more receptive to schema-inconsistent data. See FISKE & TAYLOR, *supra* note 26, at 124. For most adults who have grown up in America, I believe racial schemas are in the former category.

⁶⁷ As a general matter, there is preferred recall for schema-relevant information. Schema-relevant does not necessarily mean schema-consistent. Whether schema-consistent or schema-inconsistent information is favored depends on numerous factors, including the stage of information processing (encoding versus retrieval); the strength of the schema (newly forming versus well-formed); and the degree of cognitive load (low versus high at the time of encoding and/or retrieval). See Fiske, *supra* note 4, at 371 (noting that when individuals are put under cognitive load, information consistent with stereotypes is more readily recalled). Finally, to complicate matters further, these various factors may have nonlinear relationships with memory recall.

A useful metaphor for a schema is a theory or hypothesis that acts as a template for perceived data. See, e.g., KUNDA, *supra* note 39, at 36–40. At the “hypothesis formation” stage, inconsistent information is favored. But during the “hypothesis validation” stage, we are looking for evidence that confirms our beliefs. Again, at the highest level of expertise and interrogation, inconsistent material might be favored. See FISKE & TAYLOR, *supra* note 26, at 128–30.

tion”), but also what we actually see and remember seeing (“visual perception”).⁶⁸

In sum, schemas automatically, efficiently, and adaptively parse the raw data pushed to our senses. These templates of categorical knowledge are applied to all entities, including human targets. Racial schemas, because they are chronically accessible, regularly influence social interactions. The obvious ways in which this might happen need little elaboration. To take an extreme example, the hate criminal will map an African American into the racial category “Black” (or some other term) by visual inspection, at which point both emotional hatred and the cognitive beliefs that fuel that hatred (for example, scapegoating) can catalyze racial violence. No fancy psychological model is necessary to make sense of such vulgar acts. The payoff of this social cognitive model comes in parsing subtler cases. For instance, what if we don’t even *see* the Black man?

3. *Automaticity.* — The Computer Crash experiment reveals that we do not have to consciously “see” the Black male face for it to influ-

⁶⁸ See Jennifer L. Eberhardt et al., *Believing Is Seeing: The Effects of Racial Labels and Implicit Beliefs on Face Perception*, 29 PERSONALITY & SOC. PSYCHOL. BULL. 360 (2003). In this study, the researchers tested whether racial labels (“White” or “Black”) would influence what participants “saw” and remembered, in terms of physical characteristics, when they encountered racially morphed faces on a computer screen. Participants were first provided demographic (including race) information about the face they would soon see. Next, they were shown for ten seconds a picture of a face created by morphing the visage of a White male and an African-American male. After a two-minute break of a different memory game, they were asked questions to test recall of demographic details (with the experimental goal of checking whether they remembered the target’s race). Then, they were shown two new faces and asked to pick which was the face they had previously seen. The truth was that *neither* was the original image. One was transformed to be more “Black,” the other to be more “White.” After two more minutes, participants were asked to sketch the target’s face (with financial incentives to be as accurate as possible).

Those participants identified as “entity theorists,” who view traits as largely immutable, *see id.* at 362, showed an assimilation effect. *See id.* at 366. In other words, depending on the demographic label they were given, they “saw” an image consistent with that label. By contrast, those participants identified as “incremental theorists,” who see traits as tentative descriptors and are drawn more to individuating information, *see id.* at 362, showed a contrast effect. *See id.* at 366. By being surprised by the disconnect between the racial label and the morphed image, incremental theorists were predicted to examine the face more closely to see the deviation from the prototype member of that category. The researchers concluded that “racial labels operated as an invisible magnet, attracting entity theorists yet simultaneously repelling incremental theorists.” *Id.* at 369.

In this brief summary, it is difficult to explore why certain participants showed an assimilation effect, whereas others showed a contrast effect. My only point here is that schemas, in this case activated through textual demographic labels, can influence what we see, not only in terms of interpretation of behavior, but also in terms of physical characteristics. *Cf.* Paul R. Wilson, *Perceptual Distortion of Height as a Function of Ascribed Academic Status*, 74 J. SOC. PSYCHOL. 97 (1968). In Wilson’s study, students were introduced to a person, variously described as a student, demonstrator, lecturer, senior lecturer, or professor. Later, when asked to guess the person’s height, the higher the social status, the taller the guessed height, ranging from 68.9 inches (student) to 72.3 inches (full professor). *Id.* at 99.

ence our behavior. The researchers in that study concluded that the Black face automatically activated schema-consistent behavior through the principle of “ideomotor action.”⁶⁹ Prior research by Patricia Devine had revealed that subliminal priming with words stereotypically associated with African Americans could lead participants to interpret ambiguous behavior as more aggressive.⁷⁰ But scientists wondered whether the result stemmed partly from using words with negative affect, such as “lazy.” The Computer Crash study answered this question by demonstrating that the mere image⁷¹ of a Black face — and a subliminal one at that — could activate a Black racial schema.

Further research has demonstrated the connection between subliminal priming (through words or pictures) and subsequent tasks, such as evaluations, interpretations, and speed tasks.⁷² These findings indicate that schemas operate not only as part of a conscious, rational deliberation that, for example, draws on racial meanings to provide base rates for Bayesian calculations (what social cognitionists might call a “controlled process”). Rather, they also operate automatically —

⁶⁹ Bargh et al., *supra* note 3, at 231; *see also id.* at 232 (“Just as the accessibility or likelihood of use of a concept increases no matter what the particular source of that accessibility, the likelihood of a behavioral response may increase from thinking about that behavior, regardless of the source of that thought.”). This idea traces back to William James, who suggested that “every representation of a movement awakens in some degree the actual movement which is its object.” *Id.* at 231 (quoting 2 WILLIAM JAMES, *THE PRINCIPLES OF PSYCHOLOGY* 526 (1890)). Bargh also notes a fascinating line of ironic control studies that suggest that people instructed not to move an object (such as a pendulum along a particular axis) or think in a particular way (sexist thoughts) did just that. *See id.* at 232. *See generally* Daniel M. Wegner, *Ironic Processes of Mental Control*, 101 *PSYCHOL. REV.* 34 (1994).

⁷⁰ *See* Patricia G. Devine, *Stereotypes and Prejudice: Their Automatic and Controlled Components*, 56 *J. PERSONALITY & SOC. PSYCHOL.* 5 (1989). In that seminal study, participants were subliminally exposed to stereotypical words associated with African Americans such as “Negroes, lazy, Blacks, blues, rhythm, Africa, stereotype, ghetto, welfare, basketball, unemployed, and plantation.” *Id.* at 10. None of these words directly mentioned aggressiveness, however. Next, participants read a passage about Donald, whose race was not identified, who engaged in the ambiguous behavior of refusing to pay rent. Those who received a heavy dose of priming (80% stereotypical words) interpreted Donald’s actions as more hostile than those who received a milder dose (20%). *See id.* at 11–12.

⁷¹ Whether the prime is a word (such as a Black or White name) or a picture (such as a face) may alter the magnitude of the implicit bias as measured by response latency. *See, e.g.*, Nilanjana Dasgupta et al., *Automatic Preference for White Americans: Eliminating the Familiarity Explanation*, 36 *J. EXPERIMENTAL SOC. PSYCHOL.* 316, 326 (2000) (suggesting that the use of names produces greater effects than the use of pictures, and speculating that pictures may be simply easier to process than words, decreasing response times across the board).

⁷² *See, e.g.*, John F. Dovidio et al., *On the Nature of Prejudice: Automatic and Controlled Processes*, 33 *J. EXPERIMENTAL SOC. PSYCHOL.* 510, 516–17 (1997) (demonstrating time differentials in classifying positive or negative words as a function of receiving subliminal flashes of Black or White faces). Studies documenting such an effect cut across social categories. For example, activating an elderly stereotype by exposure to words associated with the elderly influenced the speed at which participants walked down the hallway to the elevator. None of the words that primed the stereotype had anything to do with physical speed. *See* Bargh et al., *supra* note 3, at 236–38.

without conscious intention and outside of our awareness (an “automatic process”).⁷³ Here we see translation of yet another critical race studies theme, that the “power of race is invisible.”⁷⁴

To summarize: we think through schemas generally, and through racial schemas specifically, which operate automatically when primed, sometimes even by subliminal stimuli. The existence of such automatic processes disturbs us because it questions our self-understanding as entirely rational, freely choosing, self-legislating actors. We are obviously not robots that mechanically respond to stimuli in precisely programmed ways. We do respond to individuating information, when we are motivated and able to do so.⁷⁵ Nevertheless, we ignore the best scientific evidence if we deny that our behavior is produced by complex superpositions of mental processes that range from the controlled, calculated, and rational to the automatic, unintended, and unnoticed. Finally, we must recognize that these biases are not random errors; rather, they have a tilt. After all, the participants in the Computer Crash experiment got more hostile, not more friendly, after being flashed Black faces. Why?

B. *Implicit Bias*

1. *The Problem: Opacity.* — Social psychologists have long sought to measure the nature and content of the racial meanings contained within our racial schemas. One way to measure is simply to ask people directly, and on such surveys, we see a substantial decline in racial stereotypes and prejudice in the past fifty years.⁷⁶ But are such self-reports trustworthy? Given prevailing social norms in favor of racial

⁷³ To be careful, we should distinguish what we mean by “automatic.” Fiske and Taylor describe five different meanings or criteria associated with the term: that the operation of some schema is unintentional (does not require any explicit goal); involuntary (always occurs in the presence of relevant environmental triggers); effortless (does not consume limited cognitive resources or processing capacity); autonomous (does not need to be controlled once the process has been initiated); and outside awareness (occurs without consciousness of initiation or of the process itself). FISKE & TAYLOR, *supra* note 26, at 271. I use the term “automatic” principally to emphasize that it is unintentional and outside our awareness. Whether these processes are truly “involuntary” will be discussed below.

⁷⁴ As my colleague Gary Blasi puts it, we cannot not play the race card. See Gary Blasi, *Advocacy Against the Stereotype: Lessons from Cognitive Social Psychology*, 49 UCLA L. REV. 1241, 1274 (2002).

⁷⁵ See generally Susan T. Fiske et al., *The Continuum Model: Ten Years Later*, in DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY 231 (Shelly Chaiken & Yaacov Trope eds., 1999) (refining the “continuum model” of schematic processing). Fiske and her colleagues also identify situations that may motivate more careful attention to individual attributes. See *id.* at 239–42.

⁷⁶ For example, in 1994 roughly 90% of White Americans said they would vote for a qualified Black man for president. In 1958, a majority answered no to the same question. Dolly Chugh, *Societal and Managerial Implications of Implicit Social Cognition: Why Milliseconds Matter*, 17 SOC. JUST. RES. 203, 206 (2004) (citing JAMES ALLEN DAVIS & TOM W. SMITH, GENERAL SOCIAL SURVEYS, 1972–1994: CUMULATIVE CODEBOOK 150 (1994)).

equality, an individual may feel awkward showing her ambivalence, anxiety, or resentment toward specific racial categories.⁷⁷

In order to fly under the political correctness radar, social psychologists and political scientists have asked indirect questions about race. Prominent examples, developed since the 1970s, include scales that measure symbolic and modern racism.⁷⁸ Unfortunately, this indirect questioning risks confounding political conservatism with racism.⁷⁹

⁷⁷ There is clear evidence of such impression management. For example, in the well-known “bogus pipeline” studies, Edward Jones and Harold Sigall convinced participants that they were attached to a machine that would measure their true attitudes regardless of what they in fact said. The participants did not know that the machine was bogus. Their explicit self-reports changed significantly when they were hooked up to this bogus machine. See Edward E. Jones & Harold Sigall, *The Bogus Pipeline: A New Paradigm for Measuring Affect and Attitude*, 76 *PSYCHOL. BULL.* 349 (1971); see also Russell H. Fazio et al., *Variability in Automatic Activation as an Unobtrusive Measure of Racial Attitudes: A Bona Fide Pipeline?*, 69 *J. PERSONALITY & SOC. PSYCHOL.* 1013, 1014 (1995) (discussing the possibility of a bona fide pipeline).

⁷⁸ “Symbolic racism,” a term introduced in the 1970s, is a belief system that combines the following ideas:

that racial discrimination is no longer a serious obstacle to blacks’ prospects for a good life; that blacks’ continuing disadvantages are due to their own unwillingness to take responsibility for their lives; and that, as a result, blacks’ continuing anger about their own treatment, their demands for better treatment, and the various kinds of special attention given to them are not truly justified.

P.J. Henry & David O. Sears, *The Symbolic Racism 2000 Scale*, 23 *POL. PSYCHOL.* 253, 254 (2002) (citation omitted); see also Donald R. Kinder & David O. Sears, *Prejudice and Politics: Symbolic Racism Versus Racial Threats to the Good Life*, 40 *J. PERSONALITY & SOC. PSYCHOL.* 414, 416 (1981).

“Modern racism” was introduced approximately one decade after “symbolic racism.” Similar in substance, the new term was chosen “to emphasize that the new racism was not the only form of prejudice to be rooted in socialization and in abstract, symbolic beliefs; old-fashioned racism had been as well.” Henry & Sears, *supra*, at 254–55; see also John B. McConahay, *Modern Racism, Ambivalence, and the Modern Racism Scale*, in *PREJUDICE, DISCRIMINATION, AND RACISM* 91, 92–93 (John F. Dovidio & Samuel L. Gaertner eds., 1986) (describing the Modern Racism Scale (MRS)).

In addition to symbolic and modern racism, other theories include ambivalent racism, see Irwin Katz et al., *Racial Ambivalence, Value Duality, and Behavior*, in *PREJUDICE, DISCRIMINATION, AND RACISM*, *supra*, at 35, 35–46, and aversive racism, see Samuel L. Gaertner & John F. Dovidio, *The Aversive Form of Racism*, in *PREJUDICE, DISCRIMINATION, AND RACISM*, *supra*, at 61, 61–62. Ambivalent racism theory suggests that the White mainstream believes that racial minorities deserve not only sympathy for victimization, but also resentment because they do not subscribe to conventional rules of hard work. See Fiske, *supra* note 4, at 359 (citing Katz et al., *supra*). Aversive racism involves outwardly endorsing egalitarianism but nonetheless having bias that we cannot quite admit to ourselves. See *id.* at 360 (citing Gaertner & Dovidio, *supra*). See generally Monica Biernat & Christian S. Crandall, *Racial Attitudes*, in *MEASURES OF POLITICAL ATTITUDES* 297 (John P. Robinson et al. eds., 1998) (compiling various measures of racial attitudes).

⁷⁹ See Fazio et al., *supra* note 77, at 1021 (questioning the usefulness of the MRS based on both confounding and reactivity concerns); see also Paul M. Sniderman & Phillip E. Tetlock, *Reflections on American Racism*, *J. SOC. ISSUES*, Summer 1986, at 173, 180–82; Paul M. Sniderman & Phillip E. Tetlock, *Symbolic Racism: Problems of Motive Attribution in Political Analysis*, *J. SOC. ISSUES*, Summer 1986, at 129, 145–47.

More troubling, we may honestly lack introspective access to the racial meanings embedded within our racial schemas.⁸⁰ Ignorance, not deception, may be the problem. Relatedly, our explicit normative and political commitments may poorly predict the cognitive processes running beneath the surface. While connected to the automaticity point, this disconnect between explicit and implicit bias raises a different issue: dissociation.⁸¹ The point here is not merely that certain mental processes will execute automatically; rather, it is that those implicit mental processes may draw on racial meanings that, upon conscious consideration, we would expressly disavow. It is as if some “Trojan Horse” virus had hijacked a portion of our brain.

2. *The Solution: Measuring Speed.* — How have social cognitionists measured the bias in racial meanings if it is so opaque? One method has been to use sequential priming procedures that take advantage of the automaticity of schemas.⁸² Such procedures begin by priming a participant with a particular stimulus, such as a word or a face, which activates a particular racial schema. Then, the racial meanings associated with that schema should alter performance on

⁸⁰ See, e.g., Allen R. McConnell & Jill M. Leibold, *Relations Among the Implicit Association Test, Discriminatory Behavior, and Explicit Measures of Racial Attitudes*, 37 J. EXPERIMENTAL SOC. PSYCHOL. 435, 435–36 (2001) (identifying limited introspective access, social desirability concerns, and experimenter effects as difficulties with explicit measures of prejudice).

⁸¹ One form of the dissociation thesis comes from Patricia Devine, who suggests that cultural knowledge is accepted early in our mental development, before critical faculties develop. This knowledge, in turn, generates the racial meanings that produce the automatic bias. In contrast, personal beliefs, which are ones we consciously endorse and think are true, develop later. See Devine, *supra* note 70, at 5–6. Thus, we see a separation between cultural stereotypes and personal stereotypes, with the former being the main variable in automatic processes.

Jody Armour has also presented a “dissociation model,” one that distinguishes between unconscious stereotypes and conscious, personally chosen prejudice. See Armour, *supra* note 50, at 744. In distinguishing between the unconscious and conscious modes, Armour and I are talking about the same phenomenon. However, Armour’s use of the terms “stereotypes” and “prejudice” — to map roughly the difference between implicit and explicit — is not standard to the social cognition literature. In particular, both implicit and explicit stereotypes (factual beliefs) may exist; the same goes for implicit and explicit prejudice (negative attitudes).

⁸² Other methods include, for example, the Linguistic Intergroup Bias (LIB) test. This test takes advantage of the fact that a particular scene in any narrative can be described at different levels of abstraction. From most concrete to most abstract, we could use descriptive action verbs, interpretive action verbs, state verbs, or adjectives. If a particular scene fits the stereotype, we tend to use abstract descriptors (he is a criminal). By contrast, if a scene runs counter to stereotype, we tend to use concrete descriptions (he stole food). See, e.g., William von Hippel et al., *The Linguistic Intergroup Bias as an Implicit Indicator of Prejudice*, 33 J. EXPERIMENTAL SOC. PSYCHOL. 490, 491 (1997). In one study, von Hippel measured the LIB of participants, then had them evaluate a videotape of a person asking for money on the street. Von Hippel found a correlation between the LIB bias and how threatening a Black requester for money was evaluated to be. *Id.* at 493, 507.

some subsequent linguistic,⁸³ interpretive,⁸⁴ or physical task.⁸⁵ If the prime and the task are schema-consistent, one expects a faster response; by contrast, if they are inconsistent, one expects a slower response.⁸⁶

The first such study was by Samuel Gaertner and John McLaughlin, who primed participants subliminally with the word “White” or “Black,” then immediately replaced the word with a string of letters that were sometimes words and sometimes gibberish.⁸⁷ Participants had to identify as quickly as possible whether the string of letters was indeed a word. The words chosen were associated with stereotypes of either Whites or Blacks. Participants were faster at recognizing a positive word such as “smart” if they had just been primed with the word “White” instead of “Black.”⁸⁸ The time differential in task execution was deemed a measure of “implicit bias.”

A flurry of studies adopted and varied this reaction-time paradigm by priming participants with some social category, subliminally or consciously (supraliminally), and then measuring whether their ability to execute some task was facilitated or hindered.⁸⁹ The Implicit Association Test (IAT) has become the state-of-the-art measurement tool.⁹⁰

⁸³ For example, researchers may task participants with deciding whether a string of letters is a word or not, selecting words that are schema-consistent or schema-inconsistent. See, e.g., Bernd Wittenbrink et al., *Evidence for Racial Prejudice at the Implicit Level and Its Relationship with Questionnaire Measures*, 72 J. PERSONALITY & SOC. PSYCHOL. 262, 265 (1997).

⁸⁴ For example, participants may be primed and then required to decide how “hostile” ambiguous behavior was. See, e.g., Devine, *supra* note 70, at 8–12.

⁸⁵ For examples of experiments containing physical tasks, see the Shooter Bias studies discussed *infra* section I.C.4, pp. 1525–28.

⁸⁶ The analogy for computer enthusiasts would be “benchmarking” — measuring how fast computer processors and peripherals execute tasks — under different conditions.

⁸⁷ Samuel L. Gaertner & John P. McLaughlin, *Racial Stereotypes: Associations and Ascriptions of Positive and Negative Characteristics*, 46 SOC. PSYCHOL. Q. 23, 23 (1983).

⁸⁸ *Id.* Although Gaertner and McLaughlin did not find the converse phenomenon, with participants responding faster to bad words when primed with the word “Black,” subsequent studies did. See, e.g., J.F. Dovidio et al., *Racial Stereotypes: The Contents of Their Cognitive Representations*, 22 J. EXPERIMENTAL SOC. PSYCHOL. 22 (1986).

⁸⁹ See, e.g., Dovidio et al., *supra* note 72, at 534 (finding “clear evidence” that the race of the face with which participants were primed — Black or White — influenced how quickly they could identify positive versus negative words). This effect occurs when using even the most abstract of social categories. For example, one study subliminally primed participants with ingroup and outgroup words such as “we” or “they.” After an ingroup prime, participants more quickly identified a positive word as “positive.” See BROWN, *supra* note 50, at 98–99.

⁹⁰ One of the attractive features of the IAT is that it generally produces much larger effects than did prior priming methodologies. See Jens B. Asendorpf et al., *Double Dissociation Between Implicit and Explicit Personality Self-Concept: The Case of Shy Behavior*, 83 J. PERSONALITY & SOC. PSYCHOL. 380, 382 (2002). For the most current recommendations on how to conduct and read IATs, see Anthony G. Greenwald et al., *Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm*, 85 J. PERSONALITY & SOC. PSYCHOL. 197 (2003). Readers are invited to take the test themselves online at the Project Implicit website. See Project Implicit, *IAT Home*, at <http://implicit.harvard.edu/implicit/demo> (last visited Feb. 13, 2005). For

The IAT examines how tightly any two concepts are associated with each other. In a typical experiment, two racial categories are compared, say “Black” and “White.” Next, two sets of stimuli (words or images) that correspond to the racial meanings (stereotypes or attitudes) associated with those categories are selected. For example, words such as “violent” and “lazy” are chosen for Blacks, and “smart” and “kind” for Whites.

Participants are shown a Black or White face and told to hit as fast as possible a key on the left or right side of the keyboard. They are also shown words stereotypically associated with Blacks or Whites and again told to hit a key on the left or right side of the keyboard. In half the runs, the Black face and Black-associated word are assigned to the *same* side of the keyboard (schema-consistent arrangement). In the other half, they are assigned *opposite* sides (schema-inconsistent arrangement). The same goes for the White face/White-associated stimulus combination.

Tasks in the schema-consistent arrangement should be easier, and so it is for most of us.⁹¹ How much easier — as measured by the time differential between the two arrangements — provides a measure of implicit bias. The obvious confounds — such as overall speed of participant’s reactions, right- or left-handedness, and familiarity with test stimuli — have been examined and shown not to undermine the IAT’s validity.⁹²

Amygdala. Millisecond measurements in front of computers may seem far removed from our commonsense understanding of “prejudice.” But consider what functional magnetic resonance imaging (fMRI) of the brain has recently shown. The amygdala is an almond-sized subcortical structure within our brains, involved in emotional

a list of other implicit bias measurement tools, see Irene V. Blair, *The Malleability of Automatic Stereotypes and Prejudice*, 6 PERSONALITY & SOC. PSYCHOL. REV. 242, 260–61 (2002).

⁹¹ To date there is no agreed-upon cognitive theory of why the implicit effect exists. When the IAT was introduced in 1998, the creators intentionally decided to provide only a minimal theory — namely that the more associated any two concepts were, the easier it would be to generate the same response. See, e.g., Anthony G. Greenwald et al., *Measuring Individual Differences in Implicit Cognition: The Implicit Association Test*, 74 J. PERSONALITY & SOC. PSYCHOL. 1464, 1469 (1998) (discussing the IAT effect only in terms of “associatively compatible” categories). With this strategy, they focused the field on producing further empirical results and validity testing of the IAT. This should not be mistaken for a neo-behavioralist approach; the IAT creators are thoroughly cognitive in their orientation. A fully specified cognitive theory of implicit bias awaits, but is high on the research agenda.

⁹² See T. Andrew Poehlman, *Understanding and Using the Implicit Association Test: III. Meta-analysis of Predictive Validity 4* (n.d.) (unpublished manuscript, on file with the Harvard Law School Library) (collecting citations). Poehlman also collects citations to studies demonstrating internal consistency and relative insensitivity to minor methodological differences, such as number of trials and target stimuli. See *id.*

learning, perceiving novel or threatening stimuli,⁹³ and fear conditioning.⁹⁴ Neuroscientists, collaborating with social cognitionists, have recently demonstrated that the amygdalas of White participants “light up” far more when they are *subliminally* shown Black faces as compared to White faces.⁹⁵ Moreover, the degree of amygdala activation is significantly correlated with participants’ IAT scores.⁹⁶ There is, however, no correlation with explicit measures of bias, which again demonstrates dissociation between explicit self-reports and implicit measures revealed by reaction-time differentials.⁹⁷

The fact that some observable behavior (differences in reaction time as measured by the IAT) maps to some neural activity (as measured by fMRI) says little: all human behavior must map to some neural activity. What is significant is that implicit bias seems connected to a particular area of the brain, the amygdala. We know that the amygdala becomes active when a person is exposed to stimuli with emotional significance, such as emotional faces, especially fearful ones.⁹⁸ While hardly definitive,⁹⁹ such neural imaging gives us greater reason to think that the IAT is measuring something real and significantly connected to emotion-laden racial mechanics.¹⁰⁰

⁹³ See Kevin N. Ochsner & Matthew D. Lieberman, *The Emergence of Social Cognitive Neuroscience*, 56 AM. PSYCHOLOGIST 717, 720 (2001).

⁹⁴ Fear conditioning is the process by which a neutral stimulus, through direct association with an aversive stimulus, acquires emotional properties. See Elizabeth A. Phelps et al., *Performance on Indirect Measures of Race Evaluation Predicts Amygdala Activation*, 12 J. COGNITIVE NEUROSCIENCE 729, 729 (2000).

⁹⁵ See William A. Cunningham et al., *Separable Neural Components in the Processing of Black and White Faces*, 15 PSYCHOL. SCI. 806, 811 (2004). This effect was not found when a similar experiment was done with famous, well-regarded Black and White faces. See Phelps et al., *supra* note 94, at 733. Still earlier research had found faster habituation by the amygdala to ingroup faces as compared to outgroup faces. See Allen J. Hart et al., *Differential Response in the Human Amygdala to Racial Outgroup vs Ingroup Face Stimuli*, 11 NEUROREPORT 2351, 2353 (2000).

⁹⁶ See Cunningham et al., *supra* note 95, at 811; see also Phelps et al., *supra* note 94, at 732 (finding “significant correlation between bias in response time on the IAT and strength of the amygdala activation”); *id.* at 733 (finding that the left-superior amygdala, which is known to be activated when exposed to fearful faces, correlated most strongly with negative implicit bias measures).

⁹⁷ See Phelps et al., *supra* note 94, at 732 (finding no correlation between the amygdala activation and the MRS).

⁹⁸ *Id.* at 730–31.

⁹⁹ This does not mean that the amygdala is necessary to generate implicit bias differentials in reaction-time measurements. See Elizabeth A. Phelps et al., *Intact Performance on an Indirect Measure of Race Bias Following Amygdala Damage*, 41 NEUROPSYCHOLOGIA 203, 203 (2003). In Phelps’s study, one subject with bilateral amygdala damage was compared to a control group of two comparable persons with intact amygdalas. No statistically significant difference in IAT scores were found. The scientists (including William Cunningham, whose more recent work is cited above) concluded that the amygdala was not absolutely necessary to generate the IAT effect. See *id.* at 206.

¹⁰⁰ There is additional evidence linking implicit cognitive processes and measurable physiological changes. Subliminally priming elderly people with a negative aging stereotype, then challeng-

3. *The Results: Pervasive Implicit Bias.* — Using the IAT and similar tools, social cognitionists have documented the existence of implicit bias against numerous social categories.¹⁰¹ According to Nilanjana Dasgupta, the “first wave” of research demonstrated that socially dominant groups have implicit bias against subordinate groups (White over non-White, for example). By her count “almost a hundred studies have documented people’s tendency to automatically associate positive characteristics with their ingroups more easily than outgroups (i.e. ingroup favoritism) as well as their tendency to associate negative characteristics with outgroups more easily than ingroups (i.e. outgroup derogation).”¹⁰² These studies address not only automatic attitudes (prejudice), but also automatic beliefs (stereotypes).¹⁰³ In the United States, bias has been found against Blacks, Latinos, Jews, Asians, non-Americans, women, gays, and the elderly.¹⁰⁴ Implicit bias against outgroups has also been found in other countries.¹⁰⁵

Fascinating is the overwhelming evidence that implicit bias measures are dissociated from explicit bias measures.¹⁰⁶ Put another way,

ing them with difficult math and verbal questions, produced higher heart rates and blood pressure. By contrast, subliminal priming with positive stereotypes decreased cardiovascular stress. Nilanjana Dasgupta, *Implicit Ingroup Favoritism, Outgroup Favoritism, and Their Behavioral Manifestations*, 17 SOC. JUST. RES. 143, 162 (2004) (describing the work of B.R. Levy).

¹⁰¹ My focus is on race, but other categories including age and gender have been explored. See, e.g., *id.* at 147 (discussing youth’s implicit bias against the elderly).

¹⁰² *Id.* at 146; see also Blair, *supra* note 90, at 242 (noting the same research trends). These results are consistent with an older set of social psychological theories called Social Identity Theory (SIT). See *infra* note 223 (describing SIT).

¹⁰³ Researchers have documented implicit bias held by Whites against Blacks, Latinos, and Asians. See Dasgupta, *supra* note 100, at 147 (cataloging studies). For example, one recent study found strong implicit bias among Whites to associate Whites as being more “American” than “Foreign” as compared to Asian Americans. See Thierry Devos & Mahzarin R. Banaji, *American = White?*, 88 J. PERSONALITY & SOC. PSYCHOL. (forthcoming Mar. 2005) (manuscript at 6–7, on file with the Harvard Law School Library). This result occurred notwithstanding the fact that researchers instructed the participants that all the photographs they were about to see were of “Americans.” See *id.* (manuscript at 12). Follow-up studies using the Go/No-go Association Test (GNAT) instrument demonstrated that this implicit bias was not driven solely by a tight association between Asian Americans and “Foreign.” See *id.* (manuscript at 16). In addition, Asian Americans were specifically excluded from the concept “American.” See *id.* For more information on the GNAT, see *infra* note 388.

¹⁰⁴ Dasgupta, *supra* note 100, at 147.

¹⁰⁵ *Id.* at 147 (citing studies regarding aborigines in Australia and Turkish immigrants in Germany).

¹⁰⁶ See, e.g., Dovidio et al., *supra* note 72, at 517 (finding no statistically significant relationship between implicit facilitation scores and the MRS). Dovidio and his colleagues conclude that “this pattern offers direct evidence that many Whites who report being nonprejudiced on traditional measures of prejudice do indeed harbor unconscious negative attitudes toward Blacks.” *Id.* at 534; see also Dasgupta et al., *supra* note 71, at 324 (finding that the correlations among five explicit measures of bias averaged $r = 0.51$, whereas the average correlation between explicit and implicit measures was much smaller, $r = 0.12$); von Hippel et al., *supra* note 82, at 507 (concluding that explicit measures of prejudice should be dissociated from implicit prejudice, as measured by the LIB test).

on a survey I may honestly self-report positive attitudes toward some social category, such as Latinos.¹⁰⁷ After all, some of my best friends are Latino. However, implicit bias tests may show that I hold negative attitudes toward that very group. This is dissociation — a discrepancy between our explicit and implicit meanings. This dissociation appears most vividly when “the group averages for conscious and unconscious measures [are] placed side-by-side using a common metric,” which reveals “wide divergences” between the means.¹⁰⁸

Web Harvest. For example, recent research by Brian Nosek, Mahzarin Banaji, and Anthony Greenwald, based on extraordinarily large data sets harvested from the Web, revealed such dissociation beautifully.¹⁰⁹ Whites exhibited some *explicit* preference for themselves over Blacks, but that explicit preference paled in comparison to their *implicit* preference.¹¹⁰ This dissociation does not necessarily indicate lack of covariation. The same research revealed a positive (al-

¹⁰⁷ Russell Fazio and his colleagues raised substantial questions about the value of explicit but indirect measurements of racial bias (specifically the MRS). See Fazio et al., *supra* note 77, at 1022 (reporting a study showing substantial reactivity of the MRS). Participants received very different scores depending on whether they believed that a White or Black experimenter would input the answers. See *id.* These results also provided further evidence of dissociation. See *id.* at 1020 (finding that MRS and implicit bias facilitation scores were not correlated in a statistically significant way at $p = 0.05$ and that the correlation was, in fact, negative).

¹⁰⁸ See Mahzarin R. Banaji, *The Opposite of a Great Truth Is Also True: Homage to Koan #7*, in PERSPECTIVISM IN SOCIAL PSYCHOLOGY: THE YIN AND YANG OF SCIENTIFIC PROGRESS 127, 131 (John T. Jost et al. eds., 2003) [hereinafter PERSPECTIVISM IN SOCIAL PSYCHOLOGY].

¹⁰⁹ See Brian A. Nosek, Mahzarin R. Banaji & Anthony G. Greenwald, *Harvesting Implicit Group Attitudes and Beliefs from a Demonstration Web Site*, 6 GROUP DYNAMICS 101, 105 (2002) (N = 192,364). Between October 1998 and April 2000, over 600,000 race-attitude tests using pictures of faces were completed on the Web. *Id.* at 102. The authors recognized that the dataset was not created through a random sampling. However, they pointed out that the Internet sample was far more diverse than the laboratory samples traditionally drawn from college students; furthermore, the massive Internet data could provide valuable comparisons with the laboratory data. See *id.* For other caveats regarding Web-based data, see *id.* at 104.

¹¹⁰ Explicit preference for Whites over Blacks measured by Cohen’s d was 0.36; by contrast, implicit preference for Whites over Blacks measured by Cohen’s d was 0.71. *Id.* at 105; see also Devos & Banaji, *supra* note 103 (manuscript at 15) (finding dissociation on the question whether African Americans were “American” or “Foreign”). In other words, on explicit measures, very little difference was found in judging Whites versus African Americans as “American.” However, on implicit measures, Whites were viewed as more “American” than “Blacks.” Subsequent experiments confirmed this basic effect, even when the conditions were purposefully designed to maximize the explicit and implicit connections between African-Americans and “American.” See *id.* (manuscript at 16–22) (finding the same implicit bias, including dissociation, with an IAT designed around photos of Black and White track-and-field stars during the Olympics).

Dissociation is not always shown between explicit and implicit beliefs or attitudes. For instance, researchers have found examples of minimal dissociation between explicit and implicit attitudes regarding homosexuality and political candidates. See Boris Egloff & Stefan C. Schmukle, *Predictive Validity of an Implicit Association Test for Assessing Anxiety*, 83 J. PERSONALITY & SOC. PSYCHOL. 1441, 1442 (2002); see also John F. Dovidio, *On the Nature of Contemporary Prejudice: The Third Wave*, 57 J. SOC. ISSUES 829, 839–40 (2001) (suggesting that more socially sensitive issues yield greater dissociation).

though often weak) relationship between an individual's implicit and explicit attitudes.¹¹¹ Those of us with the greatest explicit bias (as measured by self-reported answers to indirect survey questions) against a racial minority tend also to have the greatest implicit bias against them, and vice versa.¹¹² To recap, although explicit and implicit biases can register at quite different levels (including entire changes in valence, from positive to negative), there is often weak covariation.

C. Behavioral Consequences

By now, even patient readers demand a payoff: Do racial schemas alter behavior? More particularly, does implicit bias represent anything besides millisecond latencies in stylized laboratory experiments? What is the evidence, for instance, that the IAT predicts any real-world behavior, much less anything that is legally actionable?¹¹³

Research addressing behavioral consequences has been called the "second wave" of implicit bias research.¹¹⁴ There is now persuasive evidence that implicit bias against a social category, as measured by instruments such as the IAT, predicts disparate behavior toward individuals mapped to that category. This occurs notwithstanding contrary explicit commitments in favor of racial equality. In other words, even if our sincere self-reports of bias score zero, we would still engage in disparate treatment of individuals on the basis of race, consistent with our racial schemas. Controlled, deliberative, rational processes are not the only forces guiding our behavior.¹¹⁵ That we are not even aware of, much less intending, such race-contingent behavior does not magically erase the harm.¹¹⁶

¹¹¹ See Banaji, *supra* note 108, at 128, 131 (identifying this paradox as a "koan"); Nosek, Banaji & Greenwald, *supra* note 109, at 111.

¹¹² See Nosek, Banaji & Greenwald, *supra* note 109, at 106 (finding that self-identified liberals had slightly lower IAT bias scores than self-identified conservatives). A correlation between explicit and implicit bias has also been found in recent work that computes a composite ethnocentrism score by measuring bias against multiple social groups. Although the correlation between explicit and implicit bias against any single group may be low, that correlation becomes more robust when a composite ethnocentrism measure is computed. See William A. Cunningham et al., *Implicit and Explicit Ethnocentrism: Revisiting the Ideologies of Prejudice*, 30 PERSONALITY & SOC. PSYCHOL. BULL. 1332, 1342 (2004) (finding substantial overlap, $r = 0.47$, between explicit and implicit ethnocentrism).

¹¹³ For a general defense of the IAT's validity, see Anthony G. Greenwald & Brian A. Nosek, *Health of the Implicit Association Test at Age 3*, 48 ZEITSCHRIFT FÜR EXPERIMENTELLE PSYCHOLOGIE 85 (2001).

¹¹⁴ See Dasgupta, *supra* note 100, at 151.

¹¹⁵ Cf. Dovidio et al., *supra* note 72, at 519–20 (suggesting that "personal attitudes . . . best predict private but controlled responses," while "implicit attitudes" best predict spontaneous behavior, and "public attitudes" best predict behavior when social desirability factors are significant).

¹¹⁶ See Ayres, *supra* note 27, at 2421 (pointing out that the harms of race-contingent behavior can be "largely independent of the decisionmaker's mens rea").

1. *Interpreting.* — Even before the rise of reaction-time measurements, social psychologists demonstrated convincingly that schemas influence interpretation.¹¹⁷ We have already discussed how activated schemas influence which stimuli we give attention to, how we encode representations of those stimuli, and how easily we retrieve information thus stored. To a first approximation, we see what we expect to see. Like well-accepted theories that guide our interpretation of data, schemas incline us to interpret data consistent with our biases.¹¹⁸

Emily or Lakisha? A recent experiment provides powerful evidence that our racial schemas, triggered simply by names, can alter how we interpret resumes. Behavioral economists Marianne Bertrand and Sendhil Mullainathan responded to over 1300 help-wanted ads¹¹⁹ in Boston and Chicago with fictitious resumes that were crafted to be comparably qualified.¹²⁰ The sole difference was that half of the resumes were randomly assigned African-American-signaling names (for example, Lakisha Washington), while the other half were assigned “White” names (for example, Emily Walsh).¹²¹ The White resumes re-

¹¹⁷ For example, in 1976 B.L. Duncan demonstrated that automatic activation of negative stereotypes can influence the interpretation of ambiguous behavior, leading people to conclude that Blacks are more hostile than Whites. Duncan generated two videos involving one student pushing another; everything was held constant between the videos except for the race of the pusher. These videos were shown to participants, all of whom were White. When the pusher was Black, 90% characterized the action as violent or aggressive; by contrast, when the pusher was White, only 40% came to the same conclusion. See Birt L. Duncan, *Differential Social Perception and Attribution of Intergroup Violence: Testing the Lower Limits of Stereotyping of Blacks*, 34 J. PERSONALITY & SOC. PSYCHOL. 590 (1976); see also BROWN, *supra* note 50, at 100 (describing Duncan’s experiment).

Andrew Sagar and Janet Schofield confirmed these results in a study of sixth-grade boys’ characterization of line drawings and accompanying verbal narratives. The “race” of the characters was set by shading in drawings that were otherwise identical. The darker the skin, the more that the ambiguous narrative, which described a bump in the hallway, was interpreted as aggressive and hostile. This bias appeared in both White and Black youths. See H. Andrew Sagar & Janet Ward Schofield, *Racial and Behavioral Cues in Black and White Children’s Perceptions of Ambiguously Aggressive Acts*, 39 J. PERSONALITY & SOC. PSYCHOL. 590, 593–95 (1980).

¹¹⁸ See, e.g., EDWARD E. SAMPSON, *DEALING WITH DIFFERENCES: AN INTRODUCTION TO THE SOCIAL PSYCHOLOGY OF PREJUDICE* 121–22 (1999) (describing studies demonstrating our tendency to seek out information confirming our hypotheses).

¹¹⁹ The jobs ranged along a large spectrum, from clerical work to cashier to management positions. MARIANNE BERTRAND & SENDHIL MULLAINATHAN, *ARE EMILY AND GREG MORE EMPLOYABLE THAN LAKISHA AND JAMAL? A FIELD EXPERIMENT ON LABOR MARKET DISCRIMINATION 2* (Nat’l Bureau of Econ. Research, Working Paper No. 9873, 2003). The experiment took place between July 2001 and May 2002. *Id.*

¹²⁰ *Id.* The researchers were careful in creating realistic resumes based on actual resumes recently posted on job search websites for the two cities within the four occupational areas of sales, administrative support, clerical services, and customer services. *Id.*

¹²¹ The experimenters generated “Black” and “White” names by analyzing frequency of names according to race data provided on birth certificates between 1974 and 1979 in Massachusetts. *Id.* at 7. Through spot surveys, any names that did not produce correct racial mappings were thrown out. *Id.*

ceived 50% more callbacks.¹²²

To study the impact of differences in resume quality, the researchers sent both standard and higher-quality resume pairs to employers.¹²³ A White higher-quality resume enjoyed a statistically significant 30% greater callback rate than the White standard resume. By contrast, a Black higher quality resume received a statistically insignificant 9% greater callback rate than the Black standard resume.¹²⁴

To explain these results, the researchers tentatively suggest “lexicographic search by employers.”¹²⁵ Given the surfeit of resumes passing their desks, employers quickly scan them, and many stop reading after seeing a Black name.¹²⁶ This phenomenon also explains why higher-quality resumes do not produce much return for African Americans — the employer never actually gets to the details.¹²⁷ Rearticulated in terms of racial schemas, the employer applies verbal rules of racial mapping to categorize applicants by name into certain racial categories. Once the names are mapped, some set of negative racial meanings (stereotypes and prejudices) are automatically activated that produce fewer callbacks for African Americans.

To be sure, this study did not measure the bias of those who reviewed the resumes either through explicit self-reports or reaction-time measurements. Instead, it simply reported behavioral consequences. Accordingly, we cannot know specifically whether the employers were conscious of or endorsed the schematic thinking that produced these results. I am, however, confident that an explicit survey given to these firms would reveal a total commitment to equal opportunity — and not entirely for impression-management reasons. In sum, although the Emily or Lakisha study suffers from some potential confounds,¹²⁸ it

¹²² *Id.* at 2–3. The authors estimated this effect to be equivalent to the benefit generated from eight years of additional experience. White names received on average one callback for every ten resumes submitted (10.08%); Black names received one callback for every fifteen submissions (6.70%). *Id.* at 10. This difference was statistically significant. *Id.* at 10 n.31. These results were uniform across industries, with no differences demonstrated for federal contractors or employers proclaiming themselves an “Equal Opportunity Employer.” *Id.* at 3. The only difference was that employers in predominantly Black neighborhoods in Chicago discriminated slightly less. *Id.*

¹²³ The higher-quality resumes featured extra labor market experience, fewer “employment holes,” employment during school, military experience, email address, honors, and computer or other special skills. Standard and higher quality resumes did not differ in terms of years of education. *Id.* at 11–12.

¹²⁴ *Id.* at 12.

¹²⁵ *Id.* at 23.

¹²⁶ *Id.*

¹²⁷ *Id.*

¹²⁸ Using names that unambiguously signal Black or White raises vexing confounds. The unambiguously Black names could have provided negative proxy information about social background (class and mainstream cultural assimilation). Of course, such racial profiling itself could be morally and legally troubling unless one could somehow argue that the proxy information ap-

speaks volumes about the continuing race discrimination in the labor market.

Agentic Backlash. So far, I have described how schemas generally influence our interpretations, but this claim is neither controversial nor novel for social psychologists. What's new is that interpretive biases have recently been correlated to reaction-time measures of implicit bias. The best experiment is in the context of gender, not race.¹²⁹ Laurie Rudman and Peter Glick examined the relationship between implicit bias against women and their job interview evaluations.¹³⁰ Four tester candidates were created for the position of a computer lab manager: agentic man, androgynous man, agentic woman, and androgynous woman. In the "agentic" profile (for both genders), the videotaped interview and "life philosophy" essay of the job candidates emphasized self-promotion and competence. In the "androgynous" profile (again, for both genders), the written essay added qualities of interdependence and cooperation.¹³¹ Half the study participants were told that they had to evaluate the candidates for a job that required masculine qualities; the other half were told that it also required some feminine qualities.¹³² After reviewing the interview tapes and the essays, participants rated the candidates on three measures: competence, social skills, and hireability.¹³³

appropriately influenced the callback decision, notwithstanding more diagnostic information available on the face of the resume itself.

To explore this possibility, the researchers collected statistics about mothers' high school graduation rates for each of the names used in the study. They found that the Black names used in the study fell below the overall Black average (61.0% versus 70.2%), while the White names chosen were above the White overall average (91.7% versus 83.9%). *Id.* at 39 tbl.11. These findings support the class-confound interpretation. That said, there was substantial heterogeneity of social backgrounds, as measured by mother's education, among the names within each race. If accurate proxy estimates were taking place, then we would expect correlations between callbacks and actual mother's education rates. In other words, "Aisha" (percentage of mothers with high school educations, 77.2%) should have received more callbacks than "Ebony" (65.6%). But all the correlations were negative, and none was statistically significant. *See id.* at 20-21, 39 tbl.11. In addition, if class were the principal driver, one would expect to see less of a callback difference between Whites and Blacks living in wealthier zip codes, but researchers saw no such results. *See id.* at 20.

¹²⁹ I recognize that what may apply to gender may not apply to race. I am confident, however, that future empirical work will demonstrate consistent findings vis-à-vis race.

¹³⁰ Laurie A. Rudman & Peter Glick, *Prescriptive Gender Stereotypes and Backlash Toward Agentic Women*, 57 J. SOC. ISSUES 743, 747-48 (2001).

¹³¹ *See id.* at 748. The videotapes featured all applicants answering questions such as "Do you like having to perform in a pressure situation?" with scripted answers such as "I tend to thrive in pressure situations." *Id.* at 749. The androgynous was added through "life philosophy" essays that either said, "Basically, there are two kinds of people, winners and losers. My goal is to be a winner" (agentic condition), or "To me, life is about being connected to other people[,] . . . help[ing] someone out" (androgynous condition). *Id.* at 750.

¹³² *See id.* at 749.

¹³³ *Id.* at 748.

The participants evaluated women differently from men in only one setting. In the feminized job condition (in which the job explicitly called for the ability to cooperate with others), the agentic female was rated less hireable than the identical agentic male.¹³⁴ The researchers isolated the mediating variable to be differences in evaluation of “social skills,” not “competence.”¹³⁵ In other words, if the job required cooperative behavior, women who showed agentic qualities were penalized more than their identical male counterparts.

In addition to rating the job applicants, the participants completed a gender IAT and explicit gender stereotype questionnaires.¹³⁶ Not surprisingly, explicit bias measures did not correlate with how participants evaluated the social skills of agentic females.¹³⁷ What did correlate were their IAT scores: the higher the implicit bias against women, the lower the social skills rating.¹³⁸

Biased interpretation can have substantial real-world consequences. Consider a teacher whose schema inclines her to set lower expectations for some students,¹³⁹ creating a self-fulfilling prophecy.¹⁴⁰ Or a grade

¹³⁴ On a 5-point scale, with 5 being the most hireable, M (agentic male) = 3.52 versus M (agentic female) = 2.84 ($p < 0.05$). See *id.* at 753. On the hireability index, there was no disparate treatment of agentic men and agentic women in the masculine job condition. Also, across both job conditions, the androgynous men and women fared equally well. To repeat, disparate treatment was seen when the job required the candidate to be “supportive,” and agentic men were not penalized but agentic women were. See *id.*

¹³⁵ *Id.* at 753–54.

¹³⁶ *Id.* at 751–52.

¹³⁷ *Id.* at 756.

¹³⁸ *Id.* ($r = -0.49$, $p < 0.001$).

¹³⁹ In another experiment, in the educational context, John Darley and Paget Gross prepared a videotape of a fourth-grade girl named Hannah solving test problems. See John M. Darley & Paget H. Gross, *A Hypothesis-Confirming Bias in Labeling Effects*, 44 J. PERSONALITY & SOC. PSYCHOL. 20, 22–23 (1983). The tape was designed to be ambiguous about her intellectual abilities. One group of participants was provided only socioeconomic status (SES) information about Hannah and asked to predict her academic abilities. Another group was given the exact same data but also this ambiguous videotape. Interestingly, those who were given just the SES data did not stereotype Hannah’s abilities; by contrast, those who saw the videotape — which was specifically designed to be ambiguous — evaluated her ability in accordance with the SES priming. See *id.* at 24–25, 27–29; see also BROWN, *supra* note 50, at 93 (summarizing Darley and Gross’s results); SAMPSON, *supra* note 118, at 121 (same). In other words, without any information besides SES status, the participants were not willing to “stereotype.” However, with a little additional data, class-based schemas inclined participants to see what they expected. Moreover, the participants could more confidently feel that they were making a judgment based on the “merits,” not some gross generalization.

¹⁴⁰ In a classic but controversial demonstration, Robert Rosenthal and Lenore Jacobson went into an American elementary school and gave the children intelligence tests. See ROBERT ROSENTHAL & LENORE JACOBSON, *PYGMALION IN THE CLASSROOM: TEACHER EXPECTATION AND PUPILS’ INTELLECTUAL DEVELOPMENT* 61, 68–69 (1968). They randomly selected 20% of the students and told only the teachers that these select children would excel in the near future. *Id.* at 70. Eight months later, the psychologists revisited the school and retested the students. Amazingly, those who had been identified *randomly* as having high potential ended up in fact overperforming. See *id.* at 68, 74–75; see also BROWN, *supra* note 50, at

school teacher who must decide who started the fight during recess. Or a jury who must decide a similar question, including the reasonableness of force and self-defense.¹⁴¹ Or students who must evaluate an outgroup teacher, especially if she has been critical of their performance.¹⁴² The Agentic Backlash study provides support for a more specific version of our tendency toward schema-consistent interpretation by demonstrating behavioral consequences of implicit bias.

Of course, schemas do not blind us entirely to individual qualities signaled by the object of interest. Students in my class, for instance, do not treat me the same as a busboy in a Korean restaurant who speaks limited English simply because we are both mapped to the category Asian. Even if race is a central schema, it is simple enough to activate the subtype of “model minority” versus the “FOB” (fresh off the boat). Moreover, my students have much higher motivation to be accurate since I am in a position of power over them. And motivation for accuracy can prevent the use of heuristics or other cognitive shortcuts when adequate cognitive resources are available.¹⁴³ Similarly, in a nighttime encounter, a Black man dressed as a police officer is not treated the same as one in a jogging suit: role schemas often dominate the field. Nonetheless, in some situations, especially in stranger-to-stranger interactions, little additional individuating information besides what we look like and how we dress is available. In such circumstances, ambiguous actions produce schema-consistent interpretations. As Susan Fiske succinctly notes, “[p]eople are hardly equal opportunity perceivers.”¹⁴⁴

2. *Performing.* — Differential assessments may not be caused entirely by subjective interpretations. Rather, racial meanings transmitted through the culture, coupled with implicit cognitive processes, may

107–08 (summarizing Rosenthal & Jacobson’s results). The explanation is that the “high potential” schema biased the teachers’ interpretations, which in turn influenced their interactions with the students in ways that ended up genuinely raising their educational performance.

There have been various criticisms about the appropriateness of this experiment as well as its methodology. Moreover, these results have not been consistently replicated in higher grades. *See id.* at 108. Nevertheless, in the educational literature, there is substantial consensus on the significant effect of teacher expectations on student performance. *See id.*; *see also* Robert Rosenthal, *Critiquing Pygmalion: A 25-Year Perspective*, 4 CURRENT DIRECTIONS PSYCHOL. SCI. 171, 172 (1995) (“[T]here are now too many new studies for even committed criticisms of disliked results to make the basic conclusion go away: Teachers’ expectations can affect pupils’ intellectual functioning.”).

¹⁴¹ *See, e.g.*, Jon Hurwitz & Mark Peffley, *Public Perceptions of Race and Crime: The Role of Racial Stereotypes*, 41 AM. J. POL. SCI. 375, 378–79 (1997) (summarizing eclectic literature demonstrating that Whites “respond more punitively to blacks than to those of their own race”).

¹⁴² *Cf.* Eberhardt et al., *supra* note 68, at 361 (citing H.H. Kelly, *The Warm-Cold Variable in First Impression of Persons*, 18 J. PERSONALITY 431 (1950)) (noting how social labels applied to teachers, such as “cold” or “warm,” influence student evaluations).

¹⁴³ *See, e.g.*, KUNDA, *supra* note 39, at 236–37.

¹⁴⁴ Fiske, *supra* note 4, at 369.

alter how we actually perform on objectively measured tests. Evidence comes from the remarkable “stereotype threat” literature launched by psychologist Claude Steele.¹⁴⁵ In a seminal experiment, Claude Steele and Joshua Aronson gave a difficult verbal test to White and Black Stanford undergraduate students. One group was informed that the test was *ability diagnostic* — testing how smart they were. Another comparable group, given the same test, was told that the test was *ability nondiagnostic* — simply a laboratory problem-solving task. In the latter condition, the Black students performed comparably to equally skilled White students.¹⁴⁶ But in the former condition, Black students greatly underperformed equally skilled White students.¹⁴⁷

The apparent explanation for this odd result is that somehow the stereotype that Blacks are intellectually inferior got activated in the former group. According to Steele, this “stereotype threat” may have raised the group’s fear that by doing poorly, they would reinforce a negative stereotype of the group they belong to. Thus, doing poorly had a “double consequence”: not only individual failure but also confirmation of the negative stereotype.¹⁴⁸ This anxiety somehow disrupted their performance. In the stereotype threat model, this “threat” does not operate by way of explicit internalization of negative self-concept; in other words, these Black students would not self-report that they are intellectually inferior to their White peers because they are Black.¹⁴⁹ In my view, the precise mechanism of performance disruption has not yet been specified. Nonetheless, the general empirical

¹⁴⁵ See generally Claude M. Steele, *A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance*, 52 AM. PSYCHOLOGIST 613 (1997).

¹⁴⁶ In these measures, “[a]nalysis of covariance was used to remove the influence of participants’ initial skills, measured by their verbal SAT scores, on their test performance.” *Id.* at 620. Some commentators have objected to this normalization in part because it obscures the raw score differential between White and Black students, which persists even in the nondiagnostic condition. See Amy L. Wax, *The Threat in the Air*, WALL ST. J., Apr. 13, 2004, at A20. A full explanation, which Wax cites, is provided by Paul R. Sackett et al., *On Interpreting Stereotype Threat as Accounting for African American–White Differences on Cognitive Tests*, 59 AM. PSYCHOLOGIST 7, 7–10 (2004). The authors write:

It is also important to note that the above observations are not meant as criticisms of Steele and Aronson’s research. Steele and Aronson clearly demonstrated a very interesting phenomenon in a series of persuasive and carefully conducted experiments. They have shown that stereotype threat can affect the performance of some students on some tests What they have not done, *and do not purport to do*, is to offer stereotype threat as the general explanation for the long-observed pattern of subgroup differences on a broad range of cognitive tests.

Id. at 10 (emphasis added).

¹⁴⁷ See Steele, *supra* note 145, at 620.

¹⁴⁸ See *id.* at 622; KUNDA, *supra* note 39, at 373–78.

¹⁴⁹ See Steele, *supra* note 145, at 617.

findings of stereotype threat have been duplicated across various social categories, including women, Latinos, and poor White students.¹⁵⁰

What is amazing is that not only can test scores be depressed, they can also be boosted.¹⁵¹ That was the finding of the Math Test study described in the Introduction.¹⁵² By unconsciously activating a particular identity, performance on difficult tests by the very same category of people could be boosted upward (Asian) or depressed downward (woman), notwithstanding the fact that the participants were already two standard deviations above the mean in SAT math scores.¹⁵³ Margaret Shih performed the same experiment in Vancouver, British Columbia, where the stereotype that Asians are good at math is less robust, as measured by surveys. In that iteration, which involved nineteen female high school students, the control group scored highest (accuracy = 59%), followed by the Asian identity group (44%), and then the female identity group (28%).¹⁵⁴ These results suggest that it is not Asian identity alone as much as the local stereotype of Asian mathematical ability that drives the differences in performance.

A later, more comprehensive study by Shih revealed further fascinating results. First, the priming cannot be blatant. When Asian Americans were asked explicit questions about Asian stereotypes and told that the test examined whether “Asian Americans are good at mathematics,” the boost disappeared.¹⁵⁵ Second, the activated stereo-

¹⁵⁰ See Margaret Shih et al., *Stereotype Performance Boosts: The Impact of Self-Relevance and the Manner of Stereotype Activation*, 83 J. PERSONALITY & SOC. PSYCHOL. 638, 638 (2002) (collecting studies).

¹⁵¹ “Stereotype boost” should be distinguished from what has been called “stereotype lift.” See Gregory M. Walton & Geoffrey L. Cohen, *Stereotype Lift*, 39 J. EXPERIMENTAL SOC. PSYCHOL. 456 (2003). “Stereotype lift is the performance boost caused by the awareness that an outgroup is negatively stereotyped.” *Id.* at 456 (emphasis omitted). Stereotype lift always involves negative stereotypes of outgroups; by contrast, stereotype boost involves positive stereotypes typically, although not always, of the ingroup. Through a meta-analysis of stereotype threat studies, Walton and Cohen compared the performance of White men in the stereotype-relevant condition (an intelligence test described as diagnostic of ability) with that of White men in the stereotype-irrelevant condition (a test characterized as free of race or ethnic biases). *Id.* at 457. They found that White men performed worse in the stereotype-irrelevant condition ($d = 0.24$; $p < 0.0001$). Translated to the scale of the SAT, there is a fifty-point lift for being a White male. *Id.* at 463.

¹⁵² See *supra* pp. 1492–93.

¹⁵³ The 46 women who participated had an average self-reported SAT math score of 750.9 ($SD = 46.53$), compared to the national average of 508 ($SD = 110$). Correcting for restriction of range, the authors found an even larger effect and greater statistical significance, at $p = 0.01$ and $r = 0.34$. See Shih et al., *supra* note 10, at 81–82.

¹⁵⁴ See *id.* at 82.

¹⁵⁵ See Shih et al., *supra* note 150, at 641, 643 (showing decreased accuracy for the supraliminally primed, but not to the point of statistical significance).

types do not have to be self-relevant.¹⁵⁶ In other words, you may not have to *be* Asian to get the boost, as long as you simply *think* Asian. Shih primed Asian and non-Asian students either subliminally or supraliminally with words associated with Asians¹⁵⁷ or with a control list. In the supraliminal runs, Asian American performance was depressed, which is consistent with the blatant priming findings.¹⁵⁸ But non-Asian-participant performance was actually boosted.¹⁵⁹ The researchers interpreted this boost according to a “perception-behavior expressway” in which activation of a trait catalyzes trait-consistent behavior.¹⁶⁰ This explanation comports with the larger theory of ideomotor action.¹⁶¹ The results changed with subliminal primes: Non-Asians failed to perform differently when the words were flashed subliminally, which suggests that they were not even sensitive to the prime, at least in the test performance context. By contrast, Asians were sensitive to the subliminal primes, perhaps because of the self-relevance of the stimuli, and they received a boost in test performance.¹⁶²

I want to be up front about the limited state of our knowledge. We have no deep understanding of such bizarre testing phenomena.¹⁶³ But even without any clear explanation, we can safely say that racial stereotypes, both negative and positive, can be activated implicitly and explicitly to alter test performance in striking ways.¹⁶⁴ We should re-

¹⁵⁶ See *id.* at 642–43; see also 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 (1998). Dijksterhuis and van Knippenberg asked individuals to spend some time imagining either professors or soccer hooligans. Guess who performed better — all in statistically significant ways — in a subsequent task of Trivial Pursuit? Those who had thought of professors. The participants were neither professors nor hooligans, but thinking of either category influenced performance. *Id.* at 870–72.

¹⁵⁷ The words included “TOKYO, HONG KONG, WONTON, SHANGHAI, KIMONO, ASIA, TAIWAN, WOK, CHINATOWN, CHANG, CHOPSTICKS, and WONG.” See Shih et al., *supra* note 150, at 642. I note that these words are distinctly East Asian, almost entirely Chinese and Japanese. I would predict different results if words associated with Filipino, South Asian, Southeast Asian, and Pacific Islander identities were included. Also, what if we altered the valence of such words, including terms such as “alien,” “accent,” and “internment”?

¹⁵⁸ *Id.* at 643.

¹⁵⁹ The boost was from a mean of 5.75 correct answers to 7.23; this result was statistically significant at $p = 0.03$. *Id.*

¹⁶⁰ *Id.* at 639 (“After being exposed to more general trait concepts, individuals will automatically infer and perform trait-consistent behavioral tendencies.”).

¹⁶¹ See *supra* note 69.

¹⁶² See Shih et al., *supra* note 150, at 643–45.

¹⁶³ Surprisingly little research has connected the stereotype threat literature to the implicit bias literature, although it seems highly likely that related cognitive processes are responsible for these phenomena.

¹⁶⁴ No serious scientist is making the claim that stereotype threat explains all testing differences; any accusation to the contrary is a strawman. The question is whether stereotype threat produces nontrivial differences in testing performance, with consequences and path dependence accumulated over a lifetime.

member stereotype threat each time we judge someone, including ourselves, on the basis of a test score.

3. *Interacting. — Nonverbal Leakage.* Recent research demonstrates that implicit bias, as measured by reaction time studies, also predicts behavior in stranger-to-stranger social interactions, such as interviews and face-to-face meetings.¹⁶⁵ Researchers have termed this phenomenon behavioral “leakage.” Allen McConnell and Jill Leibold were the first to demonstrate the linkage between IAT results and intergroup behavior.¹⁶⁶ In this study, White participants completed an explicit bias survey and took the IAT. They were guided through the first part of the experiment by a White female experimenter but through the last part of the experiment by a Black female experimenter.¹⁶⁷ Both experimenters asked questions of participants according to a prepared script. Participants’ interactions with both experimenters were videotaped.

Trained judges blind to the participants’ bias scores coded the videotaped interactions, focusing on nonverbal behaviors such as friendliness, eye contact, and number of speech errors.¹⁶⁸ In addition, the experimenters were asked to evaluate their interactions with each participant. A strong correlation was found between the IAT scores and the ratings of both the judges as well as the experimenters:¹⁶⁹ “[L]arger IAT effect scores predicted greater speaking time, more smiling, more extemporaneous social comments, fewer speech errors, and fewer speech hesitations in interactions with the White (vs[.] Black) experimenter.”¹⁷⁰

¹⁶⁵ See Fazio et al., *supra* note 77, at 1026.

¹⁶⁶ The authors recognized that the behavior link had been demonstrated for other implicit bias measures. See McConnell & Leibold, *supra* note 80, at 436; see also Fazio et al., *supra* note 77, at 1018 (showing correlation between “facilitation” scores and subjective ratings of Black experimenter regarding friendliness and interest of participant).

¹⁶⁷ See McConnell & Leibold, *supra* note 80, at 437–38. One critic of the experimental procedure has pointed out that the events of the experiment were not counterbalanced. In particular, the topic of the first conversation (with the White experimenter) was on a neutral topic, whereas the topic of the second (with the Black experimenter) was about race and occurred after the IAT. See Chugh, *supra* note 76, at 212.

¹⁶⁸ Other observed behaviors included abruptness, general comfort level, degree of laughter, forward body lean, direction of body facing experimenter, openness of arms, expressiveness of arms, distance between seats, speaking time, number of smiles, number of speech hesitations, number of fidgeting body movements, and number of extemporaneous social comments. McConnell & Leibold, *supra* note 80, at 438.

¹⁶⁹ See *id.* at 439.

¹⁷⁰ *Id.* Dovidio had performed a similar study earlier. After undergoing an implicit bias evaluation, White participants were required to interact with a Black experimenter. This interaction was videotaped and evaluated by third parties. The White participants’ evaluations of the Black interviewer were correlated with explicit, rather than implicit, measures of bias. By contrast, nonverbal behaviors (as measured by visual contact and blinking) correlated with implicit measures, not explicit self-reports. See Dovidio et al., *supra* note 72, at 529–30.

These nonverbal behaviors that leak out from our implicit bias¹⁷¹ influence the quality of our social interactions. In classic experiments by Carl Word, Mark Zanna, and Joel Cooper, White interviewers were trained to display less friendly nonverbal behavior — the sort that has now been correlated with higher implicit bias against racial minorities.¹⁷² When such behavior was performed in front of naïve White interviewees, those interviewees gave objectively worse interviews, as measured by third parties blind to the purpose of the experiment.¹⁷³ In addition, the perceiver's (interviewer's) unfriendly nonverbal behavior can instigate retaliatory responses from the target (interviewee), causing a positive feedback loop. This creates a vicious circle that reinforces the racial schema. Worse, the perceiver's decision not to hire the target based on that social interaction is understood as legitimately on "the merits."¹⁷⁴

This phenomenon was demonstrated nicely in an experiment by Mark Chen and John Bargh.¹⁷⁵ Participants played a "password" game, in which one player (guesser) had to guess a word based on clues provided by a partner (clue-giver). All participants were White. The clue-givers were subliminally primed with either a White or Black face. No guesser in either group received a racial prime.¹⁷⁶

As frustration increased throughout the game, different levels of hostility and aggression were expected. Those clue-givers primed with Black faces were expected to show greater hostility. If there was a self-fulfilling prophecy, in which the clue-giver's hostility catalyzed the

¹⁷¹ Nonverbal behaviors leak out through other implicit associations as well. Experiments have demonstrated that certain implicit self-concepts of personality (as measured by the IAT) can better predict our nonverbal behavior than explicit self-concepts, as provided by self-reports. *See, e.g.,* Asendorpf et al., *supra* note 90, at 387 (employing a double dissociation model to find that the IAT better predicted nonverbal, spontaneous shy behavior, whereas self-reports of shyness better predicted controlled behaviors); Egloff & Schmukle, *supra* note 110, at 1449 (confirming predictive validity of IAT-based anxiety measures).

¹⁷² *See* Carl O. Word et al., *The Nonverbal Mediation of Self-Fulfilling Prophecies in Interracial Interaction*, 10 J. EXPERIMENTAL SOC. PSYCHOL. 109 (1974).

¹⁷³ *See id.*

¹⁷⁴ *See, e.g.,* KUNDA, *supra* note 39, at 349 ("[W]e may often believe that our reactions to a stereotyped individual are free of prejudice because they are based on the individual's behavior and attributes rather than on the stereotype. What we may not realize is that the very meaning of these behaviors and attributes has been colored by the stereotype."); *see also* Dovidio, *supra* note 110. Dovidio points out that in an interracial interaction, a White person who is aware of only his controlled behavior may think he treated the Black person in a warm and friendly manner. However, the Black person experiences not only this controlled behavior but also the spontaneous behavior correlated with implicit bias, and the net behavioral phenomenon may be negative. *Id.* at 842. This disjunction imparts a *Rashomon* quality to interactions in which participants exhibit high levels of dissociation.

¹⁷⁵ *See* Mark Chen & John A. Bargh, *Nonconscious Behavioral Confirmation Processes: The Self-Fulfilling Consequences of Automatic Stereotype Activation*, 33 J. EXPERIMENTAL SOC. PSYCHOL. 541, 554–55 (1997).

¹⁷⁶ *See id.* at 548–51 (describing experimental method).

guesser's hostility, then we would expect to see greater hostility on both sides (not just on the part of the clue-giver). The audio of the game was recorded and responses by clue-giver and guesser were randomly shuffled, then evaluated by third parties unaware of who received what prime. The Black-face-primed clue-givers *and their partner guessers* were evaluated as more hostile — providing evidence of the vicious circle.¹⁷⁷

4. *Shooting.* — But for some of us, things get much, much worse. Recall the Shooter Bias study. Under threat conditions that police officers face, our racial schemas incline us to shoot Black men faster. Keith Payne performed the first gun study in 2001.¹⁷⁸ He subliminally primed non-Black participants with a Black or White face and subsequently asked them to identify as fast as possible whether the object displayed was a tool or gun.¹⁷⁹ Those who had been primed with the Black face were quicker to identify guns correctly. By contrast, those primed with the White face more quickly identified tools correctly.¹⁸⁰ When participants were time-pressured to force more errors, those primed with a Black face erred more in mistaking a tool for a gun (false alarm).¹⁸¹ In this study, various forms of dissociation appeared again.¹⁸²

Joshua Correll and his colleagues performed a second gun study in 2002.¹⁸³ They created a simple videogame that placed White or Black targets holding either guns or other objects (such as wallets, soft drinks, or cell phones) into realistic background settings. There was no prior priming with a face; the target's face would act as a simultaneous prime. The researchers directed participants to decide as soon as possible whether to shoot or not shoot. In the main study, which employed almost all White participants,¹⁸⁴ the experiment revealed that participants more quickly came to the correct decision to shoot armed targets when the target was Black. Conversely, participants more quickly came to the correct decision not to shoot unarmed targets when the target was White.¹⁸⁵ In a variation of this study, again with

¹⁷⁷ See *id.* at 554–55. Further, one could not easily justify the hostility as somehow “warranted” due to an accurate stereotype of the target: everyone playing the game was White. *Id.* at 555.

¹⁷⁸ See Payne, *supra* note 15, at 183–86.

¹⁷⁹ See *id.* at 184.

¹⁸⁰ *Id.* at 185.

¹⁸¹ *Id.* at 188.

¹⁸² See *id.* at 187, 189–90.

¹⁸³ See *supra* p. 1493.

¹⁸⁴ The participants were twenty-four female and sixteen male undergraduate students at the University of Colorado at Boulder. One male was Latino. Correll et al., *supra* note 14, at 1315.

¹⁸⁵ *Id.* at 1317.

nearly all White participants,¹⁸⁶ the researchers forced errors by decreasing the amount of time available to respond.¹⁸⁷ Consistent with Payne's earlier results, participants were more likely to trigger "false alarms" against a Black target (that is, shoot when no gun was present);¹⁸⁸ conversely, they were more likely to "miss" against a White target (that is, not shoot when a gun was present).¹⁸⁹

The researchers next tested whether "shooter bias" (as measured by the difference in response times to White and Black targets) was correlated with other bias measures. In the next experimental iteration,¹⁹⁰ in addition to playing the videogame, participants answered a battery of questions that included numerous indirect tests for racism.¹⁹¹ They were also asked about their personal views of the violence, dangerousness, and aggressiveness of African Americans (an explicit measure of a *personal stereotype*, reflecting actual endorsement of the stereotype). Finally, they were asked how most White Americans would answer the same question (an explicit measure of a *cultural stereotype*, reflecting mere knowledge of the stereotype). The personal stereotype measure, reflecting endorsement, showed no correlation with shooter bias — again, demonstrating dissociation. Interestingly, what did correlate was the measure of the cultural stereotype. The more stereotypes that participants thought that *other* Whites had against Blacks, the greater their shooter bias.¹⁹² Finally, experimenters recruited Black participants to play the game.¹⁹³ They obtained similar results: the race of

¹⁸⁶ In the second study, forty-four students participated (thirty-three female and eleven male). One male was Latino; one female was Asian. *Id.* at 1318.

¹⁸⁷ The financial incentives of the game were also changed to increase the participants' desire both to be accurate and to respond within the allotted time (630 milliseconds). *See id.*

¹⁸⁸ This finding was statistically significant at $p < 0.02$. *Id.* at 1319; *see also supra* note 16.

¹⁸⁹ Correll et al., *supra* note 14, at 1319.

¹⁹⁰ In the third study, forty-eight undergraduates participated (twenty-six female and twenty-two male). Nearly all were White; two males were Latino and one female was Asian. Three females — one African-American and two White — were excluded from the analysis for background and testing-related reasons. *Id.* at 1321.

¹⁹¹ These tests included the MRS, the Discrimination and Diversity Scales, the Motivation to Control Prejudiced Responding Scale, and a few questions from the Right-Wing Authoritarianism Scale and the Personal Need for Structure Scale for good measure. *See id.*

¹⁹² Since cultural stereotype measures were not correlated with personal stereotype measures, one could not say that the former was a simple proxy for the latter. A weak correlation was found between these two measures for the subset of participants who scored low in motivation to control prejudice. But even regressing out this confounding variable, so as to focus solely on cultural stereotypes, the correlation between the cultural stereotype measure and shooter bias was still statistically significant. *See id.* at 1323.

¹⁹³ These players had to be drawn from a different pool: with insufficient numbers of Black students on campus, the researchers had to scour bus stations, malls, and food courts in Denver, Colorado, to play the game on laptops. Fifty-two adults played the game on laptops. Twenty-five of them were African American; six female and nineteen male. The sample also included twenty-one Whites (eight female, thirteen male), one Asian, four Hispanics or Latinos, and one partici-

the player surprisingly had no impact on shooter bias.¹⁹⁴

Charles Judd and his colleagues performed a follow-up study in 2004 to identify what types of racial meanings generate the shooter bias¹⁹⁵ — negative emotional affect (negatively valenced evaluations of Blacks), cognitive stereotype (linking Blacks to guns), or some combination (stereotyping associated with a particular evaluative valence).¹⁹⁶ Participants¹⁹⁷ were primed with a Black or White face.¹⁹⁸ The subsequent task involved categorizing a photograph as a handgun or insect. While both categories are negatively valenced, only the first category is stereotypically associated with Blacks. Researchers next asked participants to categorize objects as either sports equipment¹⁹⁹ or fruits. Both categories are positively valenced, but only the first category is stereotypically associated with Blacks. If (negative) prejudice were the sole source of the shooter bias, then we would expect to see no facilitation in categorizing sports equipment after a Black prime. By contrast, if stereotypes were the sole cause, then we would

pant who did not self-identify. These last three categories were removed from the analysis, though this exclusion did not affect any results. *See id.* at 1324.

¹⁹⁴ *See id.* The only other variable with statistically significant correlation with shooter bias was, unbelievably, the amount of contact that the participants claimed to have with African Americans. *See id.* at 1325. Before rushing to the conclusion that interracial contact leads to increased implicit bias, we should note that the “contact hypothesis” requires interaction under specific conditions in order to decrease racial prejudice, as conventionally measured and defined. *See* Norman Miller & Marilyn B. Brewer, *The Social Psychology of Desegregation: An Introduction, in* GROUPS IN CONTACT: THE PSYCHOLOGY OF DESEGREGATION 1, 2 (Norman Miller & Marilyn B. Brewer eds., 1984) (identifying critical requirements as exposure to disconfirming data, interaction among people of equal status, cooperation, nonsuperficial contact, and equality norms). For a recent study that finds correlations between interracial friendships and lower implicit bias, see Christopher L. Aberson et al., *Implicit Bias and Contact: The Role of Interethnic Friendships*, 144 J. SOC. PSYCHOL. 335 (2004). Aberson asked White participants for self-reports of the number of close outgroup friends: African Americans in one experiment and Latinos in another. On the basis of these self-reports, he classified participants into two categories: “no friends” or “friends.” Then he ran standard IATs. Those with outgroup friends had lower implicit bias scores. *See id.* at 340, 343 (African Americans and Latinos, respectively). This finding reflects only correlation, not causation. Consistent with general dissociation findings, the friendship measure had no correlations with measures of explicit bias. *See id.* at 341, 344 (African Americans and Latinos, respectively). One concern with this study is the use of self-reports of “close friends,” apparently without much guidance or definition of the term.

¹⁹⁵ Charles M. Judd et al., *Automatic Stereotypes vs. Automatic Prejudice: Sorting Out the Possibilities in the Payne (2001) Weapon Paradigm*, 40 J. EXPERIMENTAL SOC. PSYCHOL. 75 (2004).

¹⁹⁶ Social cognitionists have observed priming triggering all three meanings. *See id.* at 75–76 (citing relevant studies).

¹⁹⁷ Of the fifty-nine final participants, twenty-four were male and thirty-five were female; none was Black. *Id.* at 77.

¹⁹⁸ The face appeared for 155 milliseconds and was thus visible to the participant (not subliminal). Participants were also instructed to pay attention to the faces, although they were told that recognizing faces was not their primary task. *Id.* at 78.

¹⁹⁹ These included basketballs, footballs, NBA jerseys, and NFL helmets. *Id.*

expect to see facilitation with both guns and sports equipment and no facilitation with insects or fruits.

Consistent with both Payne and Correll, the experimenters discovered that participants categorized guns faster when primed with a Black face.²⁰⁰ They also found, however, faster categorization of sports equipment when primed with a Black face.²⁰¹ The effect was slightly larger with the gun, but not in a statistically significant way. Accordingly, the researchers concluded that stereotypes, rather than prejudice, best explain the shooter bias results.²⁰² If borne out in further studies, there would be practical consequences in terms of solution strategies. For example, to decrease racially disparate shooting outcomes, the task would not be to look for cops who have positive as opposed to negative evaluations of Blacks; rather, the goal would be to sever somehow the widely shared stereotypical linkage between the Black male body and guns.²⁰³

Recall Amadou Diallo, the young West African immigrant standing in the doorway to his apartment, who was shot at forty-one times by New York police who “saw” a gun that did not exist.²⁰⁴ It should haunt us to read social science that suggests that if Diallo were White, he may still be alive. For those who doubt race played any such role, the Shooter Bias studies cannot be pooh-poohed as another tiresome play of the “race card.” For those who always knew race mattered, here is cold quantification. And more chilling is the fact that Whites and Blacks both exhibited shooter bias — a contention that would be hard to make politically without the test results.²⁰⁵

D. Objections

1. *Correction Is Easy.* — Can’t individuals of good will compensate for so-called implicit biases by force of will? In other words, if my interpretations might be biased, I can will myself to be more accurate and seek out individuating information. If my performance might be altered by stereotype threat, I can will myself to concentrate and not be “psyched out.” If I might treat a racial minority differently in an interview, I can will myself to ask friendly questions, to smile, and not

²⁰⁰ *Id.* at 78–79.

²⁰¹ *Id.*

²⁰² *See id.* at 80.

²⁰³ *See id.* at 81. One could, of course, argue that the linkage is deserved “on the merits” since the stereotype is somehow accurate. I address these arguments *infra* section II.C.3.

²⁰⁴ *See, e.g.,* Correll et al., *supra* note 14, at 1314.

²⁰⁵ *Cf.* James D. Johnson et al., *Converging Interracial Consequences of Exposure to Violent Rap Music on Stereotypical Attributions of Blacks*, 36 J. EXPERIMENTAL SOC. PSYCHOL. 233, 245 (2000) (confirming that both Black and White participants were similarly influenced in their judgments of a Black target’s hostility and intelligence by the prime of a violent rap song, performed by a Black artist).

to cross my arms during the interaction. If I might shoot minorities more quickly, I can simply will myself not to do so; that is the moral thing to do. In neurobiological terms, if my amygdala gets excited, my prefrontal cortex can be commanded to compensate moments later.²⁰⁶ This is the “correction is easy” objection.

On the one hand, this challenge raises an important point: just because behavior is automatic does not mean that it is immutable. On the other hand, this challenge underappreciates how difficult such behavior may be to avoid: simply “willing” a more correct or unbiased result may produce little benefit.

First, in order to counter otherwise automatic behavior, one must accept the existence of the problem in the first place. In other words, we must be both aware of the bias and motivated to counter it. If we instead trust our own explicit self-reports about bias — namely, that we have none — we will have no motivation to self-correct. One point of this Article is to alert readers of a dissociated implicit bias and its automatic consequences.

Second, even if such will (predicated on awareness of the potential problem) exists, success is not guaranteed because compensation may require cognitive resources that are unavailable.²⁰⁷ For example, when a police officer must decide to shoot or not shoot under extraordinary pressure, the time necessary to compute the correct response does not exist. One might think that a high motivation to be unbiased, especially in life-and-death situations, could compensate notwithstanding time pressures. But another sobering study by Keith Payne demonstrated that participants explicitly instructed to *avoid using race* ironically performed worse (although not in a statistically significant way) than participants told nothing at all.²⁰⁸ These limitations are consis-

²⁰⁶ For example, in the amygdala study, *supra* note 95, there was no differential activation in the supraliminal exposure condition, when face primes were presented for 525 milliseconds. That said, different portions of the brain (the dorsolateral prefrontal cortex, ventrolateral prefrontal cortex, and the anterior cingulate), associated by neuroscientists with “regulation and executive function,” showed greater activation. Apparently, the motivation not to be prejudiced, which according to self-reports all the participants had, allowed these portions of the brain to repress the amygdala when given the time and the cognitive resources to do so. In fact, the greater the attitudinal conflict the participant exhibited, the greater the neural activity in the ventrolateral prefrontal cortex area, which is associated with attitudinal ambivalence. See Cunningham et al., *supra* note 95, at 810–11.

²⁰⁷ See Patricia G. Devine & Margo J. Monteith, *Automaticity and Control in Stereotyping*, in *DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY*, *supra* note 75, at 339, 346–47 (identifying three requirements for overriding automatic response based on stereotype activation: awareness of stereotypes, cognitive capacity, and motivation to counter).

²⁰⁸ B. Keith Payne et al., *Best Laid Plans: Effects of Goals on Accessibility Bias and Cognitive Control in Race-Based Misperceptions of Weapons*, 38 J. EXPERIMENTAL SOC. PSYCHOL. 384, 390–91 (2002). In fact, the “avoid race” group performed only marginally worse than another group told explicitly to “use race”; that difference was also not statistically significant. *Id.* at 390. Applying a technique called “process dissociation,” the authors concluded that the racial primes

tent with findings that even when we are told what the IAT measures, we cannot through simple conscious “will” erase the time differentials:²⁰⁹ we cannot be politically correct even when it is embarrassing not to be. Also, certain behaviors, such as eye-blinking and nervousness, are simply less controllable through conscious will. There is no evidence, for instance, that the test-disrupting anxiety identified as “stereotype threat” can be mitigated by repeating the mantra “I will not be psyched out.”

By no means am I suggesting that self-correction is impossible. An accuracy motivation — fueled by the carrot of financial reward or the stick of accountability — coupled with adequate cognitive resources will correct against many automatic biases.²¹⁰ For instance, if all employers in the Emily or Lakisha study were told that an independent third party would judge their screening interview decisions for fairness and that a finding of any racial preference toward Whites (or Blacks) would lead to the death penalty, we would likely have found no statistically significant bias against (or for) African Americans. In fact, in some cases, we might find individuals overcompensating for what they fear to be the behavioral consequences of their implicit bias.²¹¹

My point is not that self-correction of bias is impossible. Rather, it is that such compensation may be difficult and, in quotidian situations, unlikely.²¹² In addition, if we are to compensate for these implicit cognitive processes, conscious, post hoc, effortful attempts at correc-

produced misperceptions of gun or tool through the automatic, not controlled, component of processing. *Id.* at 391–93. By telling participants to avoid using race, the experimenters had simply made the racial schema more accessible for the automatic processing. *Id.* at 395.

²⁰⁹ See, e.g., McConnell & Leibold, *supra* note 80, at 436 (suggesting that the IAT cannot be “gamed,” because even those who know what the IAT is supposed to measure reliably produce the IAT effect). Actually, although we cannot easily game the IAT by trying to speed up our schema-inconsistent reactions, we can purposefully slow down our schema-consistent reactions to decrease the reaction time differential. See Do-Yeong Kim, *Voluntary Controllability of the Implicit Association Test (IAT)*, 66 SOC. PSYCHOL. Q. 83, 91 (2003). This slow-down strategy should be fairly easy to detect, however.

²¹⁰ See generally Serena Chen & Shelly Chaiken, *The Heuristic-Systematic Model in Its Broader Context*, in DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY, *supra* note 75, at 73, 76–79 (discussing the interrelations between heuristic and systematic models of processing, and identifying multiple motivations such as accuracy, self-interest, and impression management). Fiske and her colleagues also discuss possible correction factors. See Fiske et al., *supra* note 75, at 239–42.

²¹¹ See Dasgupta, *supra* note 100, at 158–59. Attempts at correction may, of course, backfire. See, e.g., *supra* notes 69, 208 (discussing various ironic processes). Consider also how conscious efforts to be “friendly” can increase awkwardness. To give one personal story, while I was a student at the Harvard Law School, a well-educated spouse of an alum found out I was Korean, then immediately shared with me how much she liked her Korean dentist. I can’t say that her well-meaning attempt improved our social interaction.

²¹² For a powerful statement of this argument, see John A. Bargh, *The Cognitive Monster: The Case Against the Controllability of Automatic Stereotype Effects*, in DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY, *supra* note 75, at 361.

tion²¹³ may not be the most effective tactic; implicit techniques that decrease implicit bias, for example, may be superior.²¹⁴

2. *Correction Is Impossible.* — Another objection raises the opposite concern: resistance is futile. Drawing on sociobiology and/or evolutionary psychology,²¹⁵ one could argue that schematic thinking simply cannot be avoided. Just as we may be hardwired to be averse to rattlesnakes and to be fond of our parents' smiles, we may simply be hardwired through hundreds of thousands of years of natural selection to dread other races and to love our own. In other words, the objection goes, race is not a social construction. It is a biological reality, to which we have inborn reactions, and almost nothing can be done about it. This is the "correction is impossible" objection.

One can accept schematic thinking as inevitable, and even concede arguing that affective reactions to certain stimuli are hardwired,²¹⁶ without abandoning the position that race is socially constructed. Recall the racial mechanics model, which specifies that the racial categories, mapping rules, and meanings associated with the categories are socially constructed. First, even if there are some basic hardwired

²¹³ Indeed, there may be a danger of "rebound," in which the more that people try to suppress stereotypes, the less successful they are in doing so. See Devine & Monteith, *supra* note 207, at 349–50.

²¹⁴ Potential techniques are outlined *infra* section II.C.2, pp. 1553–63. For a general account of the difference between explicit motivated individuation and cognitive inhibition, see Galen V. Bodenhausen et al., *On the Dialectics of Discrimination: Dual Processes in Social Stereotyping*, in DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY, *supra* note 75, at 271, 284 (explaining difference between explicit motivated individuation and cognitive inhibition). See also Bargh, *supra* note 212, at 377–78 (discussing possibility that automatic egalitarian processes might counter automatic stereotyping); Eliot R. Smith & Jamie DeCoster, *Associative and Rule-Based Processing: A Connectionist Interpretation of Dual-Process Models*, in DUAL-PROCESS THEORIES IN SOCIAL PSYCHOLOGY, *supra* note 75, at 323, 329–30 (distinguishing associative processing — what I call implicit processes — from rule-based processing — what I call explicit processes); cf. Sung Hui Kim, *The Banality of Fraud: Resituating the Inside Counsel as Gatekeeper*, 74 FORDHAM L. REV. (forthcoming 2005) (manuscript at 5, on file with the Harvard Law School Library) (similarly arguing that altering the ethical situation through structural rearrangements can stop fraud better than explicit moral commandments can).

²¹⁵ In the law reviews, see the work of Owen Jones. See, e.g., Owen D. Jones, *Law and Biology: Toward an Integrated Model of Human Behavior*, 8 J. CONTEMP. LEGAL ISSUES 167 (1997); Owen D. Jones, *Time-Shifted Rationality and the Law of Law's Leverage: Behavioral Economics Meets Behavioral Biology*, 95 NW. U. L. REV. 1141 (2001). My comments in this section are not a broad indictment of behavioral biology, especially when such work is carefully pursued. Fruitful work remains to be done in connecting this literature with the implicit bias research cited in this paper.

²¹⁶ Mark Chen and John Bargh have provided evidence that we have an automatic reaction to all stimuli in terms of positive or negative affect, which has consequences on our behavioral disposition either to approach or avoid any given object. See Mark Chen & John A. Bargh, *Consequences of Automatic Evaluation: Immediate Behavioral Predispositions To Approach or Avoid the Stimulus*, 25 PERSONALITY & SOC. PSYCHOL. BULL. 215, 218–23 (1999). There is some evidence for this view, for instance, in the similar colorism found in otherwise disparate cultures. See Dasgupta, *supra* note 100, at 149–50.

mapping rules that use gross morphology to classify human beings into affective clumps, it is a logical leap to say that race — as legally and culturally experienced in the United States in the twenty-first century — is hardwired. According to current racial categories and mapping rules, fair-skinned Northern Chinese are put into the same racial category as Filipinos and Asian Indians; Blancos and Morenos are both “Latinos”; visually similar American Indians and Koreans are not in the same racial category. There is no evidence, however, that our brains are hardwired precisely in this manner.

Second, even if we are hardwired to clump people by some phenotype (regardless whether they align identically with current racial categories and their mapping rules), the meanings associated with those categories are not immutably set. As evidence, consider how racial meanings can change radically within a lifetime. At the end of the nineteenth century, the illegal immigration problem in America had a Chinese face. The Chinese were viewed as inscrutable, subhuman, incapable of higher learning; useful laborers but otherwise despicable; vectors for disease, filth, and immorality.²¹⁷ And now, the racial meaning ascribed to the very same body is often “model minority.”²¹⁸ To be sure, some meanings — such as unfair competitor and forever foreign — have persisted over time.²¹⁹ But it would be disingenuous to deny substantial transformations in both the cognitive and affective content toward Asian people.²²⁰ While explicit and implicit biases against that category have by no means disappeared, they have transformed within one lifetime. Natural selection simply does not work this quickly.

Finally, the “correction is impossible” objection could suggest that because implicit bias in favor of one’s ingroup is inevitable, it should not be morally chastised. But as an empirical matter, all groups do not show such ingroup bias. Recall the Web Harvest study, which provided evidence of pervasive implicit bias among Whites, dissociated from their explicit self-reports of racial equality.²²¹ In that study, the researchers also examined racial minorities and their schematic baggage. They inquired, for instance, whether belonging to a minority category somehow immunized members from implicit bias against

²¹⁷ See *supra* note 55.

²¹⁸ For a review of changing socioeconomics and demographics that led up to the “model minority” stereotype, see Eric K. Yamamoto, Margaret Chon, Carol L. Izumi, Jerry Kang & Frank H. Wu, RACE, RIGHTS, AND REPARATION: LAW AND THE JAPANESE AMERICAN INTERNMENT 258–71 (2001).

²¹⁹ I discussed these racial meanings in my student note in the Harvard Law Review. See Note, *supra* note 55, at 1930–33.

²²⁰ For a similar point about the potentially rapid rate of change of racial meanings as applied to Latinos, see Haney López, *supra* note 52, at 33 & n.125.

²²¹ See *supra* pp. 1513–14.

their own category.²²² On the one hand, according to theories such as Social Identity Theory (SIT),²²³ we would expect to see the same ingroup favoritism and outgroup derogation seen in Whites. This prediction would also be consistent with an evolutionary psychology story that speculates that ingroup bias is adaptive. On the other hand, we might see an opposite impulse. John Jost and Mahzarin Banaji's System Justification Theory (SJT),²²⁴ for instance, recognizes our need to feel that the world is fair and just. Therefore, large disparities in wealth, power, and social success as accreted through history are interpreted as "warranted" on the merits.²²⁵ Racial categories low on the

²²² Nosek, Banaji & Greenwald, *supra* note 109, at 111–12. In a follow-up study to Devine's seminal experiment, *see* Devine, *supra* note 70, Russell Fazio and his colleagues demonstrated that African Americans and Whites have different automatic reactions to pictures of African Americans; whereas Whites had relatively negative automatic reactions, African Americans had relatively positive ones. *See* Fazio et al., *supra* note 77, at 1019. The researchers first showed participants either a Black or White face. They then showed participants a word and asked them to classify that word as either good or bad. *See id.* at 1016. White participants classified negative words faster after Black primes, and positive words faster after White primes. *See id.* at 1017. However, for Black participants, the researchers found the opposite effect. *See id.* Moreover, in both groups they noticed substantial variability among individuals, suggesting that a "personal evaluation," not merely some "shared cultural stereotype," was being activated. *See id.* at 1019.

²²³ Crudely summarized, SIT contends that if we want to feel better about ourselves, we must feel better about the groups with which we identify and to which we belong. In the process, we must create a devalued outgroup. Meanings associated with both ingroup and outgroup are tilted accordingly. This phenomenon is well demonstrated, for example, by the work of Henri Tajfel and his "Minimal Group Paradigm." *See* BROWN, *supra* note 50, at 45–47 (describing experiments performed by Tajfel in 1971 and summarizing findings). Tajfel classified participants into two groups based on their expressed preferences for two rather similar looking paintings, one by Kandinsky and the other by Klee. They were told nothing else about other participants, except the group to which they were assigned (Kandinsky preferer or Klee preferer). Participants were then instructed to select monetary payouts for both ingroup and outgroup members from a given table of options. *See id.* at 46.

Notwithstanding the weakness of the group identity, 70% favored their ingroup and gave them a higher dollar payout. *Id.* at 46–47. In certain iterations of the experiment, some participants were willing to sacrifice absolute benefit for comparative advantage (choosing \$12 for ingroup and \$11 for outgroup, over \$13 for both groups) based on the thinnest of group identities — the preference for one Kandinsky over another Klee. *Id.* Subsequent experiments within this paradigm have demonstrated a "remarkably robust phenomenon" replicated over twenty times in multiple countries. *Id.* at 47.

²²⁴ *See* Alice H. Eagly & Amanda Diekmann, *The Commonsense Psychology of Changing Social Groups*, in PERSPECTIVISM IN SOCIAL PSYCHOLOGY, *supra* note 108, at 203–04; *see also* John T. Jost et al., *Non-Conscious Forms of System Justification: Implicit and Behavioral Preferences for Higher Status Groups*, 38 J. EXPERIMENTAL SOC. PSYCHOL. 586, 589–97 (2002) (providing evidence of outgroup favoritism by low-status groups — of Stanford University by San Jose State University students, of Whites by Latinos and Asian Americans, of men by women — through implicit measures, not explicit self-reports).

²²⁵ *See* Devos & Banaji, *supra* note 103, at 5 (identifying "history of intergroup relations," "actual hierarchy of social groups," and "internalization of that hierarchy" as possible factors).

hierarchy are thus seen as deserving their lowly status — even if one belongs to such a group.²²⁶

The data reveal that racial minorities show *both* impulses: to favor not only their ingroup (consistent with SIT), but also those on top of the racial hierarchy — Whites (consistent with SJT). According to Nosek, Banaji, and Greenwald's processing of over 17,000 runs of the IAT, African Americans on average exhibited no implicit ingroup favoritism; in fact, depending on how the data set was circumscribed, they showed a slight bias in favor of Whites or no bias either way.²²⁷ This is the mean; the distribution around the mean looks like a bell curve. By contrast, as already stated, Whites show substantial implicit ingroup favoritism on the IAT.²²⁸ If we expected only ingroup favoritism to be in play, then African Americans should have shown substantial favoritism for Blacks (at the expense of Whites); after all, on *explicit* measures, Blacks showed even greater ingroup favoritism for themselves than Whites did for themselves.²²⁹ But on implicit measures, on average, African Americans exhibited no favoritism.²³⁰ These results have been replicated for other racial groups as well, suggesting that implicit ingroup bias is not inevitable.²³¹

²²⁶ Cf. Stangor & Schaller, *supra* note 39, at 21–22 (suggesting that stereotypes rationalize the status quo).

²²⁷ See Nosek, Banaji & Greenwald, *supra* note 109, at 105–06.

²²⁸ In terms of Cohen's *d*, the IAT effects in favor of Whites were: Whites (0.83), Asians (0.78), Hispanics (0.66), and Blacks (0.16). *Id.* at 110.

²²⁹ Again, in terms of Cohen's *d*, explicit Black ingroup preference (0.80) was much higher than explicit White ingroup preference (0.59). These results are consistent with laboratory experimental findings. *Id.* at 105–06.

²³⁰ See *id.*; see also Robert W. Livingston, *The Role of Perceived Negativity in the Moderation of African Americans' Implicit and Explicit Racial Attitudes*, 38 J. EXPERIMENTAL SOC. PSYCHOL. 405, 411 (2002) (finding similar results in a separate study).

²³¹ See Eric Uhlmann et al., *Subgroup Prejudice Based on Skin Color Among Hispanics in the United States and Latin America*, 20 SOC. COGNITION 198, 202 (2002) (listing studies). Uhlmann and his colleagues ran IATs comparing Latinos and Whites and found no ingroup favoritism among Latinos in the United States. See *id.* at 220 (proposing Jost and Banaji's SJT as explanation). This result obtained regardless whether the face prime was light-skinned (Blanco) or dark-skinned (Moreno). See *id.* at 206. Researchers had reason, however, to think that skin color would matter since previous studies had found implicit bias against Morenos within the Latino group. See *id.* at 201–02. Among self-identified Blancos, there was clear implicit preference for Blancos; even self-identified Morenos favored Blancos, although to a lesser degree. See *id.* at 206. None of this bias manifested itself in explicit surveys. See *id.* at 211.

To test the significance of place and cultural context, the experimenters conducted a similar study in Chile. *Id.* at 211–12. Latinos in Chile responded similarly to Latinos in the United States. They too implicitly favored Whites, regardless of the use of a Blanco or Moreno prime. See *id.* at 213–14. As for intraracial colorism, Chileans again responded similarly. Those who self-identified as Blancos demonstrated an implicit preference for Blancos (over Morenos) that was stronger than that found within the United States. See *id.* at 213. The same bias in favor of Blancos existed for self-identified Morenos, although the number of such participants was quite small. See *id.* What was different was that Chileans were comfortable displaying *explicit* preference for Blancos as well (thus not demonstrating dissociation). See *id.* at 215. In Latin America,

Also, as a normative matter, we must not conflate “is” and “ought.” Even if it is descriptively true that we are hardwired to have implicit bias in favor of our “race” (or clumps of people loosely affiliated with today’s social construction of race), that says nothing about what we should do about it normatively. If resistance were truly futile, one could question moral disapprobation of what is genetically determined. But we know that this is not the case.

E. A Research Agenda

My model of racial mechanics is a simple application of schematic thinking. We map individuals to racial categories according to the prevailing racial mapping rules, which in turn activates racial meanings that alter our interaction with those individuals. The mapping and activation are automatic, and the racial meanings that influence our interaction may be stereotypes and prejudice we explicitly disavow. But disavowal does not mean disappearance, and it turns out that reaction time measures, such as the IAT, can measure the latent persistence of these implicit racial meanings. And implicit bias has behavioral consequences, which can be deadly.

Shooter Bias,²³² while most graphic, should not overshadow the cumulative effects of more banal encounters. Because of background stereotypes, you may do worse than you would have otherwise. See Math Test.²³³ Those objective scores may put you on par with another applicant, but because of the racial schemas triggered by your name, you may never get the interview. See Emily or Lakisha.²³⁴ And even when you do get the interview, Agentic Backlash²³⁵ suggests that you may be interpreted as having worse skills because of the social category to which you are mapped. In addition, the interview may go badly because of Nonverbal Leakage.²³⁶ The total impact of these interactional phenomena on education (admissions, mentoring), employ-

there appears to be less stigma attached to preferring Blancos over Morenos. An explanation may be an ideology of racial democracy that conceptualizes such preferences as merely aesthetic, not racial or status-related and not especially invidious. *See id.* at 200, 215.

Another study found that Asian Americans also had implicit bias against Asian Americans on the American/Foreign IAT, notwithstanding that participants were told that all pictures they saw in the test were of “Americans.” *See* Devos & Banaji, *supra* note 103 (manuscript at 29).

²³² *See supra* p. 1493.

²³³ *See supra* pp. 1492–93.

²³⁴ *See supra* pp. 1515–17.

²³⁵ *See supra* pp. 1517–19.

²³⁶ *See supra* section I.C.3, pp. 1523–25; *see also* Chugh, *supra* note 76, at 211–18 (describing the behavioral consequences for managers not only in terms of their interpersonal role, but also in terms of their informational and decisionmaking roles). Even if you get in the door, implicit bias may get in the way of teamwork. *See, e.g.,* Dovidio, *supra* note 110, at 844–45 (reporting data that White/Black teams with “aversive racists” — defined as Whites with low explicit but high implicit bias against Blacks — were the least efficient in completing a cooperative task).

ment (hiring, promotion), social networking (friendship, marriage, collegiality), and market transactions (auto purchases, mortgages) cannot be underestimated.

As future research confirms, constrains, and elaborates these results, a vast research agenda will open for those who explore the nexus of law and racial mechanics. Topics on that agenda include:

- the role of intent in all bodies of law;²³⁷
- criminal law (for example, racial profiling, self-defense, community policing, jury selection,²³⁸ penalty setting²³⁹);
- antidiscrimination law (for example, disparate treatment,²⁴⁰ disparate impact, unconscious discrimination,²⁴¹ hostile environments, mortgage lending);

²³⁷ See, e.g., Reshma M. Saujani, *"The Implicit Association Test": A Measure of Unconscious Racism in Legislative Decision-Making*, 8 MICH. J. RACE & L. 395, 413–15 (2003) (suggesting IAT scores as an additional evidentiary factor for measuring racial intent under the rule of *Village of Arlington Heights v. Metropolitan History Development Corp.*, 429 U.S. 252, 267–68 (1977)).

²³⁸ See, e.g., *id.* at 419–20 (discussing the use of IAT in jury selection process). Consider also what SJT might have to say about the significance of minority representation on juries. The Supreme Court, in *Castaneda v. Partida*, 430 U.S. 482 (1977), cautioned that “[b]ecause of the many facets of human motivation, it would be unwise to presume as a matter of law that human beings of one definable group will not discriminate against other members of their group.” *Id.* at 499 (addressing discrimination against Mexican Americans in grand jury selection); see also *id.* at 503 (Marshall, J., concurring) (pointing to social scientists’ agreement that “members of minority groups frequently respond to discrimination and prejudice by attempting to disassociate themselves from the group, even to the point of adopting the majority’s negative attitudes towards the minority”).

²³⁹ See, e.g., Irene V. Blair et al., *The Influence of Afrocentric Facial Features in Criminal Sentencing*, 15 PSYCHOL. SCI. 674, 677 (2004) (finding no disparate sentencing on the basis of race in Florida data set, but finding that within each racial category, White or Black, those individuals with more afrocentric facial features received harsher sentences); see also *United States v. Clary*, 34 F.3d 709, 710 (8th Cir. 1994) (rejecting the district court’s ruling that the crack cocaine statute and U.S. Sentencing Guidelines violated African-American defendants’ equal protection rights). The Eighth Circuit specifically rejected the lower court’s reliance on “unconscious racism” and disputed the idea that stereotypical media representations of crack users influenced Congress. *Id.* at 713. Legal interest in the findings of social cognitionists may lead to a reopening of the issue. Cf. Theodore Eisenberg & Sheri Lynn Johnson, *Implicit Racial Attitudes of Death Penalty Lawyers*, 53 DEPAUL L. REV. 1539, 1553 (2004) (finding that capital defense attorneys have implicit bias scores similar to the rest of the population).

²⁴⁰ See, e.g., Linda Hamilton Krieger, *The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity*, 47 STAN. L. REV. 1161, 1164–65 (1995) (arguing that current Title VII jurisprudence does not adequately address unconscious bias); Rebecca Hanner White & Linda Hamilton Krieger, *Whose Motive Matters?: Discrimination in Multi-Actor Employment Decision Making*, 61 LA. L. REV. 495, 499 (2001) (arguing that disparate treatment inquiries in employment discrimination cases “should focus on causation, not conscious discrimination”).

²⁴¹ Obviously, this issue has been on the research agenda for critical race theory for a long time. See, e.g., Lawrence, *supra* note 28, at 324–26 (focusing on the problem of diffuse, unconscious, collective racism).

- civil rights law and policy (for example, affirmative action's²⁴² contact hypothesis,²⁴³ role model justifications, merit definitions, advocacy strategies, housing segregation);
- lawyering and evidence (for example, strategies and rules with which to engage jurors' implicit biases);²⁴⁴
- education law and policy (for example, teaching strategies, interpretation of tests, debiasing programs and environments);
- privacy law (for example, comparing measures of implicit bias, such as the IAT, with polygraph results; widespread use of fMRI brain scans; IATs for Article III confirmations or legislators²⁴⁵);
- labor law (for example, comparing IATs to other psychological tests, such as the Myers-Briggs test, given before hiring or promotion; employment discrimination; new compliance intermediaries; evidentiary privileges for voluntary debiasing programs²⁴⁶);
- constitutional law (for example, equal protection intent versus impact, autonomy as a constitutional value, paternalism);
- cultural policy (for example, spectrum regulation, campus speech codes, subsidization of production and distribution of debiasing content, media ownership policy);
- remedies, both voluntary and court-ordered (for example, requiring debiasing screensavers as part of a settlement in a discrimination suit; providing debiasing booths in lobbies where jurors wait to be picked; providing debiasing software installed on computers).

²⁴² See, e.g., Linda Hamilton Krieger, *Civil Rights Perestroika: Intergroup Relations After Affirmative Action*, 86 CAL. L. REV. 1251, 1257–58 (1998) (discussing the relevance of social cognition and social identity theory to affirmative action policy).

²⁴³ See, e.g., Aberson, *supra* note 194, at 344 (finding correlations between close friendships with outgroup minorities and lower implicit bias scores).

²⁴⁴ See, e.g., Armour, *supra* note 50, at 766–72 (arguing against formal colorblindness and in favor of allowing counsel to address issues of race directly in order to counteract potential bias among jurors).

²⁴⁵ See, e.g., Saujani, *supra* note 237, at 414 (suggesting that legislators could be forced to take the IAT “on the stand”). Leading social cognitionists, such as Anthony Greenwald and Mahzarin Banaji, would resist such crude applications of the IAT.

²⁴⁶ See, e.g., Deana A. Pollard, *Unconscious Bias and Self-Critical Analysis: The Case for a Qualified Evidentiary Equal Employment Opportunity Privilege*, 74 WASH. L. REV. 913, 916 (1999) (persuasively calling for “the recognition of a privilege for unconscious-bias testing to encourage its use in equal employment opportunity efforts”).

If we move beyond law and consider markets, social norms, and architecture,²⁴⁷ the research possibilities become myriad.²⁴⁸ This paper serves as an explicit call for legal scholars to join the investigation. We can follow the lead of Linda Hamilton Krieger,²⁴⁹ Gary Blasi,²⁵⁰ and Jody Armour,²⁵¹ who have already blazed significant paths.

Some might say that I am calling for an overeager extension of a premature science, embraced for political reasons. And one must concede that science has been and will always be exploited for political purposes.²⁵² Just as the Right might jump on *Bell Curve*²⁵³ findings, the Left might jump on stereotype threat findings. There will always be those who out of convenience declare faith in some set of scientific explanations without due diligence. Accordingly, the goal has to be honest, public, and transparent engagement on the merits.

This requires, for instance, highlighting scientific findings that cut against one's political orthodoxy. The most vivid example this Article points out is the fact that even African Americans seem to suffer from shooter bias. I also point out that Asian Americans generally have implicit biases against African Americans that are almost as strong as those held by Whites. Neither finding is convenient to progressive politics, but that does not mean they should be swept under the rug. And in this Article, they are not.

Finally, in demanding due diligence of scientific models and explanations, we must not privilege the status quo's conventional wisdom as somehow apolitical and uncontroversial. For example, right now the dominant scientific description of human behavior within legal discourse is Rational Choice Theory.²⁵⁴ Even though there is compelling

²⁴⁷ This approach comes from cybersocial theory, whose styles of thinking will become more relevant in Part II. The standard legal scholarship citations are to LAWRENCE LESSIG, *CODE AND OTHER LAWS OF CYBERSPACE* (1999); and Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, 76 *TEX. L. REV.* 553 (1998).

²⁴⁸ Imagine, for example, designing video games and virtual world experiences that decrease implicit bias through repeated exposure to the appropriate stimuli. See *infra* pp. 1571-72 (discussing video games); cf. Kang, *supra* note 19, at 1160-79 (discussing the design of "integration spaces" for online communities in accordance with the contact hypothesis literature).

²⁴⁹ See sources cited *supra* notes 240, 242.

²⁵⁰ See Blasi, *supra* note 74.

²⁵¹ See Armour, *supra* note 50; see also JODY DAVID ARMOUR, *NEGROPHOBIA AND REASONABLE RACISM: THE HIDDEN COSTS OF BEING BLACK IN AMERICA* (1997).

²⁵² Cf. Oscar H. Gandy, Jr., *Journalists and Academics and the Delivery of Race Statistics: Being a Statistician Means Never Having To Say You're Certain*, 4 *RACE & SOC'Y* 149 (2001) (describing the problematic use of racial statistics by journalists, public intellectuals, and expert witnesses).

²⁵³ RICHARD J. HERRNSTEIN & CHARLES MURRAY, *THE BELL CURVE* (1994).

²⁵⁴ See Russell Korobkin, *Problems with Heuristics for Law*, in *HEURISTICS AND THE LAW* (Gerd Gigerenzer & Christopher Engle eds., forthcoming 2005) (manuscript at 1-2) ("All versions of [Rational Choice Theory] assume that actors will process information, make choices, and exe-

evidence that this model thoroughly misdescribes human behavior, it is still deployed as the dominant model, from which a laundry list of minor deviations are conceded.²⁵⁵ For those who complain that acceptance of the implicit bias science is “politically” motivated, one can make the same claim of the status quo.²⁵⁶

Recognizing our self-understandings to be provisional, we must still confront the difficult choices to come. As social cognitionists further demonstrate the possibility of altering levels of implicit bias — and explore the mechanisms to do so most efficiently — we will encounter difficult philosophical and legal questions about our autonomy, our normative commitments to racial equality, and the proper role of explicit collective action by private and public actors to decrease implicit bias.²⁵⁷

II. TROJAN HORSES

A. *Tuning In to Broadcast*

In the second half of this Article, I pursue a concrete application of the racial mechanics model. This Part concerns, of all things, recent FCC decisions about the local news. To understand my choice of topic, we must start with a fundamental question: “Where do racial meanings come from?” Racial meanings that accrete in our schemas can, on the one hand, come from “direct experiences” with individuals mapped into those categories. On the other hand, the racial meanings can arise from what I call “vicarious experiences,” which are stories of or simulated engagements with racial others provided through various forms of the media or narrated by parents and our peers.²⁵⁸ Given

cute behaviors in a way calculated to maximize their expected utility — that is, maximize the differential between expected benefits and expected costs.”).

²⁵⁵ See Tanina Rostain, *Educating Homo Economicus: Cautionary Notes on the New Behavioral Law and Economics Movement*, 34 LAW & SOC’Y REV. 973, 974–75 (2000) (pointing out how both behavioral law and economics theorists and social norms theorists “routinely insist that they are not engaged in a fundamental critique of law and economics”).

²⁵⁶ Jon Hanson and David Yosifon observe that science is often “captured” by politics. 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, 206–12 (2003) (describing the lesson of Galileo and his recanting).

²⁵⁷ Those uninterested in mass media policy or FCC regulation may be ready to stop reading. As explained earlier, the two Parts were written to be read in a modular fashion. However, I encourage readers to at least review the “malleability” studies in section II.C.2, *infra* pp. 1553–63, so as to avoid an unnecessarily pessimistic view of the scientific research.

²⁵⁸ See BROWN, *supra* note 50, at 83 (describing the sociocultural view that stereotypes are embedded in culture, and conveyed and reproduced through socialization and “repeated exposure to images in books, television and newspapers”); *cf.* FISKE & TAYLOR, *supra* note 26, at 147 (noting that schemas develop either from actual encounters or from “abstracted communications”).

persistent racial segregation, we should not underestimate the significance of vicarious experiences.²⁵⁹ Even if direct experience with racial minorities more powerfully shapes our schemas,²⁶⁰ vicarious experiences may well dominate in terms of sheer quantity and frequency.²⁶¹

The next question becomes, “Why are racial meanings biased against racial minorities?” One hypothesis is that people encounter skewed data sets — or as the computer scientists say, “garbage in, garbage out.”²⁶² If these principally vicarious experiences, transmitted through electronic media, are somehow “skewed,” then the racial meanings associated with certain racial categories should also be skewed.²⁶³ This analysis invites further study of culture and mass media policy, topics that social cognitionists have largely avoided.²⁶⁴

Suppose that social cognitionists identify which types of vicarious experiences trigger and exacerbate bias and which ameliorate it. Private parties will obviously be free to act on the basis of such discoveries. Voluntary attempts to create a “diversity” of role models on television reflect some such impulse, in addition to financial self-interest since “diversity” is sometimes good for business. But what about collective action, mediated through the state and implemented through law?

One could argue that there is a third source: evolutionary biology. For my views on the claim that racial meanings are provided by natural selection, see *supra* section I.D.2, pp. 1531–35.

²⁵⁹ See Kang, *supra* note 19, at 1166–69; Kang, *supra* note 36, at 1163.

²⁶⁰ See, e.g., FISKE & TAYLOR, *supra* note 26, at 520–21 (suggesting that attitudes formed through direct experience are more specific, confidently held, stable, and resistant to counterargument as compared to attitudes based on indirect experience).

²⁶¹ See ROBERT M. ENTMAN & ANDREW ROJECKI, *THE BLACK IMAGE IN THE WHITE MIND* 49 (2000) (suggesting that what I call “vicarious experiences” dominates in the construction of mainstream culture). One can view policies of affirmative action, public education, and residential desegregation partly as attempts to increase direct experiences with racial others to counter the impact of skewed vicarious experiences. This is the essence of the social contact hypothesis.

²⁶² As discussed above, groups may have a self-serving preference for the ingroup, regardless of skewed datasets. See *supra* pp. 1512–14. But the fact that subordinated groups have less implicit bias in favor of themselves than dominant groups have for themselves reveals that sociocultural hierarchy matters.

²⁶³ See Armour, *supra* note 50, at 755–56 (emphasizing the significance of mass media reactivation of stereotypes); *id.* at 764 (suggesting that “an omnipresent mass media and its incessant manipulation of stereotypes” have entrenched anti-Black stereotypes as deeply today as during Clarence Darrow’s time).

²⁶⁴ See, e.g., AUGOUSTINOS & WALKER, *supra* note 39, at 3 (suggesting that social psychology is dominated by an individualistic orientation that does not pay enough attention to “social” influences); Stangor & Schaller, *supra* note 39, at 6 (criticizing the social cognitive approach for underappreciating societal influences, including indirect sources of information such as the media). For a comprehensive analysis of race and communications, see OSCAR H. GANDY, JR., *COMMUNICATION AND RACE: A STRUCTURAL PERSPECTIVE* (1998).

Maybe the state can do nothing. Recall *American Booksellers Ass'n v. Hudnut*.²⁶⁵ Persuaded by the feminist scholarship of Catharine MacKinnon, Indianapolis enacted a city ordinance criminalizing pornography. The regulation was immediately challenged on First Amendment grounds. The Seventh Circuit, per Judge Frank Easterbrook, conceded the full empirical foundation for the ordinance for purposes of argument. The court accepted that pornography was not so much about rational persuasion as about socialization, which influences attitudes.²⁶⁶ The court further accepted that pornography indeed perpetuates gender subordination.²⁶⁷ Nonetheless, pornography could not be banned;²⁶⁸ to do so, as Easterbrook put it, would be “thought control.”²⁶⁹

The fact that pornography does not persuade through reasoned argument and instead functions through more implicit vectors was of no legal significance. As Easterbrook pointed out, the same could be said of Hitler’s orations.²⁷⁰ He further explained:

Sexual responses often are unthinking responses, and the association of sexual arousal with the subordination of women therefore may have a substantial effect. But almost all cultural stimuli provoke unconscious responses. Religious ceremonies condition their participants. Teachers convey messages by selecting what not to cover; the implicit message about what is off limits or unthinkable may be more powerful than the messages for which they present rational argument. Television scripts contain unarticulated assumptions. People may be conditioned in subtle ways. If the fact that speech plays a role in a process of conditioning were enough to permit governmental regulation, that would be the end of freedom of speech.²⁷¹

The stridency of Judge Easterbrook’s remarks seems to suggest that the state’s attempts to regulate media images are doomed to fail. But there is one communications medium that has always tolerated substantial state intervention: broadcast. In the United States, broadcast is regulated in a public-private partnership. As the Communications Act of 1934²⁷² makes clear, the electromagnetic spectrum that broadcasters employ as the wireless “channel” of communications is not private property. Instead, it is owned by the government, held in public

²⁶⁵ 771 F.2d 323 (7th Cir. 1985), *aff'd*, 475 U.S. 1001 (1986).

²⁶⁶ *See id.* at 328–29.

²⁶⁷ *See id.* at 329.

²⁶⁸ *See id.* at 332.

²⁶⁹ *Id.* at 328. Judge Easterbrook continued, noting that the ordinance “establishes an ‘approved’ view of women, of how they may react to sexual encounters, of how the sexes may relate to each other. Those who espouse the approved view may use sexual images; those who do not, may not.” *Id.*

²⁷⁰ *See id.* at 329.

²⁷¹ *Id.* at 330.

²⁷² 47 U.S.C. §§ 151–615 (2000).

trust for all.²⁷³ The United States licenses that spectrum to private parties who exploit that resource not only for private commercial gain but also for the “public interest.”²⁷⁴ No one may broadcast without a license from the federal government, which authorizes the use of a particular frequency, at a specified transmission power, within a designated geographical area. A broadcast license is limited in time — currently lasting eight years²⁷⁵ — although renewal is nearly certain.

In the 1934 Act, Congress created the FCC and charged it with managing the spectrum to further the “public convenience, interest, or necessity”²⁷⁶ — the public interest standard. In addition to regulating entry by assigning frequencies,²⁷⁷ the FCC has power to mold, at least softly, the content of broadcast. Given our robust constitutional and political commitment to free expression, one might wonder how such

²⁷³ See *id.* § 301.

²⁷⁴ See, e.g., *id.* §§ 307, 309. The “public interest” standard has been much criticized as a vague, shifting standard that has not done much to actually promote the public’s interest. See, e.g., NEWTON N. MINOW & CRAIG L. LAMAY, *ABANDONED IN THE WASTELAND: CHILDREN, TELEVISION, AND THE FIRST AMENDMENT* 58–104 (1995) (finding no concrete benefits for children so far); Reed E. Hundt, *The Public’s Airwaves: What Does the Public Interest Require of Television Broadcasters?*, 45 DUKE L.J. 1089, 1098 (1996) (calling for the standard to be applied with greater clarity and administrability); Erwin G. Krasnow & Jack N. Goodman, *The “Public Interest” Standard: The Search for the Holy Grail*, 50 FED. COMM. L.J. 605, 607–08 (1998) (finding the standard to be vague and changing); Randolph J. May, *The Public Interest Standard: Is It Too Indeterminate To Be Constitutional?*, 53 FED. COMM. L.J. 427, 428–29 (2001) (suggesting that the standard violates the nondelegation doctrine). Not surprisingly, many commentators have called for abandoning traditional regulation based on the public interest doctrine. See, e.g., Daniel L. Brenner, *Ownership and Content Regulation in Merging and Emerging Media*, 45 DEPAUL L. REV. 1009, 1012 (1996) (arguing that antitrust regulation alone will suffice); Robert Corn-Revere, *Red Lion and the Culture of Regulation*, in *RATIONALES & RATIONALIZATIONS: REGULATING THE ELECTRONIC MEDIA* 1, 12–13 (Robert Corn-Revere ed., 1997) (rejecting the scarcity rationale underlying the government’s ability to regulate media in the public interest); Gregory L. Rosston & Jeffrey S. Steinberg, *Using Market-Based Spectrum Policy To Promote the Public Interest*, 50 FED. COMM. L.J. 87, 115 (1997) (calling for more market-based approaches).

Ronald Krotoszynski, for one, suggests that there are three alternatives to the public-interest model of regulation: a system of spectrum fees or auctions, the imposition of common carriage requirements, and a pure market-based approach. He favors the first approach as the most efficient way to subsidize programming that the market would not otherwise provide. See Ronald J. Krotoszynski, Jr., *The Inevitable Wasteland: Why the Public Trustee Model of Broadcast Television Regulation Must Fail*, 95 MICH. L. REV. 2101, 2126 (1997) (reviewing MINOW & LAMAY, *supra*); see also Henry Geller, *Public Interest Regulation in the Digital TV Era*, 16 CARDOZO ARTS & ENT. L.J. 341, 362 (1998).

My criticism of the recent rearticulation of the public interest standard should not be framed as a rejection of broadcast exceptionalism. On this point, I find Jonathan Weinberg’s thoughts insightful. See Jonathan Weinberg, *Broadcasting and Speech*, 81 CAL. L. REV. 1101, 1109–10, 1199–1200 (1993) (recognizing two conflicting worldviews — privileged and nonprivileged — in First Amendment jurisprudence, with broadcast doctrine representing an enclave where nonprivileged positions are manifested).

²⁷⁵ See 47 U.S.C. § 307(c)(1) (2000); see also *id.* § 309(h) (making clear the limited license rights that are given to applicants).

²⁷⁶ See *id.* § 303.

²⁷⁷ See *id.* § 303(c).

constraints are tolerated.²⁷⁸ But under current First Amendment law, the medium matters: the Supreme Court has accepted scarcity²⁷⁹ and intrusiveness/unique availability²⁸⁰ justifications to permit greater regulation of the spectrum, as compared to other media, such as print.²⁸¹ Deeper theoretical justifications for this exceptionalism have been proffered, if not expressly adopted, by the Court.²⁸²

In its history, the FCC has promulgated (and the courts have enforced) regulations that restrict the broadcast of content deemed “bad,” such as obscenity,²⁸³ indecency,²⁸⁴ and excessive commercialization.²⁸⁵ Specific to antiracism, the FCC, at the instruction of the courts, has revoked the broadcast licenses of stations that favored segregation and aired anti-Black racial epithets.²⁸⁶ Conversely, the FCC has also promulgated regulations that promote content deemed “good” through informational programming guidelines,²⁸⁷ community needs and inter-

²⁷⁸ Cf. *id.* § 326 (“Nothing in this chapter shall be understood or construed to give the Commission the power of censorship over the radio communications or signals transmitted by any radio station, and no regulation or condition shall be promulgated or fixed by the Commission which shall interfere with the right of free speech by means of radio communication.”). Since the Radio Act of 1927, this provision has been read to prohibit only censorship of programming in advance of broadcast. See *FCC v. Pacifica Found.*, 438 U.S. 726, 735–38 (1978).

²⁷⁹ See *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 390–92 (1969).

²⁸⁰ See *Pacifica*, 438 U.S. at 729–30, 748–49 (citing the “uniquely pervasive presence” of broadcast, which is “uniquely accessible to children,” as grounds for allowing the FCC to put a negative mark in a broadcast station’s license file for broadcasting comedian George Carlin’s indecent “Filthy Words” monologue).

²⁸¹ See *Miami Herald Publ’g Co. v. Tornillo*, 418 U.S. 241, 243–44, 258 (1974) (striking down a Florida right-of-reply statute, as applied to a newspaper).

²⁸² See Lee C. Bollinger, Jr., *Freedom of the Press and Public Access: Toward a Theory of Partial Regulation of the Mass Media*, 75 MICH. L. REV. 1, 1–3 (1976); Weinberg, *supra* note 274, at 1171.

²⁸³ See, e.g., 18 U.S.C. § 1464 (2000) (criminalizing radio communication of “obscene, indecent, or profane” content); 47 C.F.R. § 73.3999(a) (2003) (prohibiting the broadcast of obscene material).

²⁸⁴ See, e.g., 18 U.S.C. § 1464; see also *Pacifica*, 438 U.S. at 748–51. In practice, the indecency ban has resulted in time channeling of potentially “indecent” speech to the safe harbor of late night. See 47 C.F.R. § 73.3999(b) (“No licensee of a radio or television broadcast station shall broadcast on any day between 6 a.m. and 10 p.m. any material which is indecent.”); *Action for Children’s Television v. FCC*, 58 F.3d 654, 656 (D.C. Cir. 1995) (upholding time channeling as constitutional).

²⁸⁵ See *The Revision of Programming and Commercialization Policies, Ascertainment Requirements, and Program Log Requirements for Commercial Television Stations*, 98 F.C.C.2d 1076, 1101 (1984) [hereinafter *TV Deregulation Order*] (describing a 1973 order setting guidelines of sixteen minutes per hour for commercials). Although most of the commercialization guidelines were lifted in 1983, the Court of Appeals for the D.C. Circuit prevented the lifting of these caps in children’s television. See *Action for Children’s Television v. FCC*, 821 F.2d 741, 744, 750 (D.C. Cir. 1987).

²⁸⁶ See, e.g., *Office of Communication of the United Church of Christ v. FCC*, 425 F.2d 543, 547–50 (D.C. Cir. 1969). See generally JERRY KANG, *COMMUNICATIONS LAW AND POLICY* 334–36 (2001) (providing background on these cases).

²⁸⁷ Until the broadcast deregulation of the early 1980s, broadcast stations were required to show at least 5% local programming, 5% informational programming (news and public affairs), or

ests ascertainment requirements,²⁸⁸ the fairness doctrine,²⁸⁹ and children's educational television guidelines.²⁹⁰ Specific to questions of race, the FCC has also tried to promote "good" and diverse content by increasing minority ownership of stations through affirmative action.²⁹¹ Finally, the FCC has regulated market structure at each stage of production, distribution, and consumption. Examples regarding production include the now-defunct Prime Time Access Rules (PTAR)²⁹² and financial-syndication ("FinSyn") rules.²⁹³ Examples regarding distribution include the various rules concerning station ownership that were altered in the recent media ownership deregulation. Examples regarding consumption include the V-chip requirement of the 1996 Telecommunications Act.²⁹⁴

10% total non-entertainment programming. See TV Deregulation Order, *supra* note 285, at 1078 (describing programming criteria established in 1976, before broadcast deregulation).

²⁸⁸ See TV Deregulation Order, *supra* note 285, at 1097 (describing then-existing ascertainment requirements).

²⁸⁹ For a general description of this now-extinct doctrine, including descriptions of the minimum threshold and responsive programming requirements, see KANG, *supra* note 286, at 511-13.

²⁹⁰ Congress passed the Children's Television Act (CTA) of 1990. See Pub. L. No. 101-437, 104 Stat. 996 (codified as amended at 47 U.S.C. §§ 303(a), 303(b), 394 (2000)). Some commentators suggest that two different public interest standards should be expressly recognized: one for adults, and another for children. See, e.g., Clay Calvert, *Toxic Television, Editorial Discretion, & the Public Interest: A Rocky Mountain Low*, 21 HASTINGS COMM. & ENT. L.J. 163, 169 (1998).

²⁹¹ See KANG, *supra* note 286, at 361-62.

²⁹² See *In re* Review of the Prime Time Access Rule, 11 F.C.C.R. 546, 547 (1996) (describing rules that generally limited network affiliates in the top fifty markets from broadcasting more than three hours of network or off-network programming during the four prime-time hours). These rules were repealed in 1995. See generally KANG, *supra* note 286, at 432-33.

²⁹³ The FinSyn rules prevented networks from syndicating network-produced programs to independent television stations. Also, networks were not allowed to purchase syndication rights from independent producers. See *Schurz Communications, Inc. v. FCC*, 982 F.2d 1043, 1045 (7th Cir. 1992). The rules were struck down in *Schurz Communications*. See *id.* at 1055. On remand, the FCC abandoned the regulations and allowed them to sunset.

²⁹⁴ In the 1996 Telecommunications Act, Congress mandated that "V-chips" be installed in all television sets "shipped in interstate commerce or manufactured in the United States" that were thirteen inches or larger along the diagonal. 47 U.S.C. § 303(x). These computer chips can read labels embedded in programming content. If a program identifies itself as having, for example, more violence than the amount set by the viewer, then that program will be blocked. Congress and the FCC were clever enough not to mandate self-labeling, which would have raised serious First Amendment concerns. Instead, the FCC "persuaded" the National Association of Broadcasters, the National Cable Television Association, and the Motion Picture Association of America to adopt a voluntary rating system. The TV rating system codifies six age levels (Y, Y7, G, PG, 14, MA) with content descriptors ("V" for violence, "S" for sexual situation, "L" for language, "D" for suggestive dialogue). Not everyone has agreed to self-label. For example, Black Entertainment Television has opted out of the program. See *In the Matter of Implementation of Section 551 of the Telecommunications Act of 1996: Video Programming Ratings*, 13 F.C.C.R. 8232, 8284-85 (1998) (separate statement of Commissioner Gloria Tristani). Also, NBC transmits age levels but not content descriptors. The rating system does not apply to sports, many movies on premium cable channels, and of course the news. This elaborate V-chip system has not been legally challenged. See KANG, *supra* note 286, at 204-07 (reviewing the V-chip system).

The point of this catalog of congressional and FCC interventions is not to defend each regulation on its merits. Further, not all of these regulations have been fully tested in court, and many may have had little impact on content. But they nonetheless show that our commitment against shaping broadcast content is far from categorical. I do not mean to conflate the descriptive with the normative; however, our lengthy history of broadcast regulation creates breathing room to speak new ideas. Given constitutional law as we know it, if we are curious about what the state can do to combat implicit bias transmitted through vicarious experiences, broadcast is the prime site of inquiry.

B. Redefining the Public Interest

The touchstone for governmental management of broadcast is the “public interest” standard. That standard has recently been explicated in an unusual way. At least in the context of ownership policy, the public interest has been functionally equated with the local news.

In June 2003, a divided FCC lifted numerous media ownership restrictions in the name of the “public interest.”²⁹⁵ Some of the changes permitted greater horizontal consolidation in local markets. Specifically, the FCC liberalized the local television multiple ownership rule²⁹⁶ and the local radio ownership rule.²⁹⁷ Some of the changes permitted greater horizontal consolidation in national markets. Specifically, the FCC altered the national television station ownership (NTSO) rule,²⁹⁸ although it chose to retain the dual network rule.²⁹⁹ Finally, some of the changes permitted greater consolidation across media markets by altering cross-ownership rules limiting radio/television³⁰⁰ and newspaper/broadcast mergers.³⁰¹ These rule

²⁹⁵ The vote was 3–2, along party lines. Media Ownership Order, *supra* note 22, at 13,620.

²⁹⁶ 47 C.F.R. § 73.3555(b) (2003).

²⁹⁷ *See id.* § 73.3555(a).

²⁹⁸ *See id.* § 73.3555(d). The NTSO rule limited the number of television stations that any single firm could own by instituting a national audience-reach cap. Such national caps had existed since 1941, and after passage of the 1996 Telecommunications Act, that cap had been set at 35%. Pub. L. No. 104-104, § 202(c), 110 Stat. 56, 111. The FCC concluded that this cap could not be justified. Media Ownership Order, *supra* note 22, at 13,818. It did not, however, entirely eliminate the cap, due to localism concerns. Without some such cap, the FCC feared that television networks would exercise too much power over their affiliates, and this concern was characterized as a concern for “localism.” *See id.* at 13,828. Accordingly, it set the national cap at 45%. *Id.* at 13,843. For a description of how this 45% cap went down to 35% and back up to 39%, see JERRY KANG, COMMUNICATIONS LAW & POLICY 63 (Supp. 2004), available at http://jerrykang.net/commlaw/kang_2004_supplement.pdf.

²⁹⁹ *See* 47 C.F.R. § 73.658(g) (2003).

³⁰⁰ *See id.* § 73.3555(c).

³⁰¹ *See id.* § 73.3555(d). Previously, broadcast/newspaper cross-ownership was generally not allowed, except for grandfathered cases and situations involving failing firms.

changes have generated tremendous controversy,³⁰² including a recent appellate court remand.³⁰³ But my purpose in recounting the order is not to explore the standard disputes.³⁰⁴ Instead, I focus on how the FCC operationalized, and thus arguably redefined, the idea of the “public interest” by equating it with the production of *local news*.

The FCC began by deconstructing “public interest”³⁰⁵ into its three constituent components: diversity, competition, and localism. As for “diversity,” the FCC recognized multiple types,³⁰⁶ but focused on viewpoint diversity — “the availability of media content reflecting a variety of perspectives” — as most important.³⁰⁷ Consistent with past practice,³⁰⁸ the Commission concerned itself only with “local, not na-

³⁰² For a brief account of the mixed congressional response to the FCC rule changes, see KANG, *supra* note 298, at 63.

³⁰³ See *Prometheus Radio Project v. FCC*, 373 F.3d 372, 430 (3d Cir. 2004) (remanding numerical limits on radio ownership to the FCC for further justification).

³⁰⁴ Accordingly, the following analysis will not critique, for example, whether any particular consolidation or cross-media regulation was liberalized too much or too little. A description of the Media Ownership Order, its controversies, and the *Prometheus* opinion is available in KANG, *supra* note 298, at 54–106.

³⁰⁵ This term has never been very clear in meaning. See Krasnow & Goodman, *supra* note 274, at 606 (suggesting that no other area in communications has generated as much controversy).

³⁰⁶ Outlet diversity was significant to the extent that there was a positive correlation between diversity in ownership of outlets and in expressed viewpoints. See Media Ownership Order, *supra* note 22, at 13,627 (noting the traditional assumption of correlation between the two); *id.* at 13,629 (affirming the “longstanding determination that the policy of limiting common ownership of multiple media outlets is the most reliable means of promoting viewpoint diversity”). The FCC, however, also recognized independent, not instrumental, advantages outlet diversity offers in terms of innovation, see *id.* at 13,632–33, and the promotion of public safety, see *id.* at 13,633.

Program diversity refers to the variety of programming formats and content, which includes, for television, the distribution of “dramas, situation comedies, reality shows, and news-magazines.” *Id.* at 13,631. For radio, program diversity refers to the variety in music formats for stations, such as “jazz, rock, and classical as well as all-sports and all-news.” *Id.* at 13,631–32. Programming targeted at “minority and ethnic groups” is also included in the definition. *Id.* at 13,632. For this sort of diversity, the FCC concluded that market forces would be the best way to provide sufficient diversity and rejected any sort of government intervention or regulation. *Id.*

Source diversity was defined as the “availability of media content from a variety of content producers.” *Id.* at 13,633. Vertical integration rules, such as PTAR and FinSyn, fall into the sort of regulations that attempt to increase source diversity. The FCC rejected source diversity as an independent aspect of diversity that should be pursued in the public interest. See *id.*

Finally, minority and female ownership diversity was confirmed as an historically significant objective. *Id.* at 13,634–35.

³⁰⁷ *Id.* at 13,627.

³⁰⁸ See, e.g., In the Matter of Amendment of Section 73.3555 of the Comm’n’s Rules Relating to Multiple Ownership, 100 F.C.C.2d 17, 23 (1984) (“Indeed, it would appear eminently reasonable to consider viewpoint diversity to be primarily a matter pertaining to local diversity, in that viewers in San Francisco, St. Louis and Philadelphia each judge viewpoint diversity by the extent of sources of ideas available to them, not by whether those same or other ideas are available in other broadcast markets.”).

tional,” diversity.³⁰⁹ Interesting was how the FCC decided to measure “viewpoint diversity”:

Although all content in visual and aural media have the potential to express viewpoints, we find that viewpoint diversity is *most easily measured* through news and public affairs programming. Not only is news programming *more easily measured* than other types of content containing viewpoints, but it relates most directly to the Commission’s core policy objective of facilitating robust democratic discourse in the media. Accordingly, we have sought in this proceeding to measure how certain ownership structures affect news output.³¹⁰

Although the FCC was willing to credit news magazine programs, such as *60 Minutes* and *Dateline*,³¹¹ it refused to consider other programming formats. The Fox Network specifically invited the FCC to credit entertainment programming that addressed or challenged stereotypes, such as “*Will & Grace*, *Ellen*, *The Cosby Show*, and *All in the Family*.”³¹² The FCC declined. Even though “entertainment programs can contribute to our goal of viewpoint diversity,” the FCC stated that it would focus instead on the news component of viewpoint diversity.³¹³ In other words, uncontroversial administrability would trump substance.

Local news also played a starring role in one other component of the public interest: “localism.” Localism has never been consistently defined in the Commission’s analysis. At times, the value of localism sounds in distributive justice terms, in the idea that rural communities and small cities should have some proportional number of broadcast licenses to big cities.³¹⁴ At other times, localism seems to focus on the degree to which broadcast content satisfies local preferences, notwithstanding national preferences or the dictates of faraway corporate

³⁰⁹ Media Ownership Order, *supra* note 22, at 13,827 (“We also affirm our decision that the market for viewpoint diversity is local, not national. And we reiterate our 1984 statement that even if the national market were the relevant area to consider, the proliferation of media outlets nationwide renders the current rule unnecessary.”); *see also id.* at 13,631 (noting that there was no reason to be concerned about national diversity given the “vast array of national news sources”).

³¹⁰ *Id.* (emphasis added).

³¹¹ It is not clear that the FCC actually did count such programs in any systematic way in its structural analyses. It instead focused on the number of hours of local news, with some discussion of news quality as measured by industry awards and/or viewer ratings.

³¹² Media Ownership Order, *supra* note 22, at 13,631.

³¹³ *Id.*; *see also id.* at 13,837 (“Other than our interest in promoting market structures that encourage local news production, we seek to avoid resting broadcast ownership policies on subjective judgments about the public policy value of different types of locally-substituted programming.”). The Order added that any aspect of viewpoint diversity contained within entertainment programming would be amply addressed by the concept of “program diversity,” which earlier the FCC said market forces could best create. *See id.* This, of course, was a non sequitur.

³¹⁴ *See id.* at 13,643 (describing initial allocation of radio licenses by the Federal Radio Commission, which insured that all communities, regardless of size, had at least one radio station, and describing this action as an early embrace of localism).

headquarters.³¹⁵ Related is the idea that national broadcast television networks should not have overweening power over local affiliates. Still other conceptions of localism emphasize coverage of local events or the production of locally produced content, perhaps using local talent, regardless whether the local community prefers such content.³¹⁶ In its order, the FCC did not clarify the term, but it did establish a methodology for measuring localism. It focused again on “programming responsive to local needs and interests, and *local news* quantity and quality.”³¹⁷ For two out of the three fundamental components of the “public interest” — diversity and localism — the FCC highlighted the significance of local news production.³¹⁸

Not surprisingly, nearly all of the structural deregulations were justified, partly or entirely, by the claim that more local news would be produced. In liberalizing local television ownership caps, for instance, the Commission explained that duopolies produced more local news through economies of scale.³¹⁹ In liberalizing the NTSO rule, the Commission concluded that network-owned stations broadcast more local news and public affairs programming than affiliates overall.³²⁰ In lifting the broadcast/newspaper ban, the commission pointed out that newspaper-owned network-affiliated stations provided almost 50% more local news and public affairs programming as compared to those network affiliates not owned by a newspaper.³²¹ Similar claims

³¹⁵ See, e.g., NEWTON N. MINOW, *The Vast Wasteland*, Address to the 39th Annual Convention of the National Association of Broadcasters (May 9, 1961), in *EQUAL TIME: THE PRIVATE BROADCASTER AND THE PUBLIC INTEREST* 48, 58 (Lawrence Laurent ed., 1964).

³¹⁶ See Media Ownership Order, *supra* note 22, at 13,644 (quoting *NBC v. United States*, 319 U.S. 190, 203 (1943), which described local program service as vital to community life and tailored to the needs of local community).

³¹⁷ *Id.* (emphasis added).

³¹⁸ The FCC recognized one final component of the “public interest”: “competition.” Nothing especially relevant to local news came up in this analysis. However, the FCC did clarify that it would measure “competition” not solely by looking at competition in advertising markets. The FCC said it would consider other metrics as well, such as measures of audience share. See *id.* at 13,639–40.

³¹⁹ See *id.* at 13,679–80 (crediting study submitted by the Fox Network); see also *id.* at 13,630 (crediting anecdotal evidence). Somewhat remarkably, the FCC said that even if jointly owned stations repeated the exact same news programs at different times, this may “actually expand the ‘amount’ of news available to viewers in that market, if viewers previously unable to watch news programming can watch the news at a different time.” *Id.* at 13,682.

³²⁰ See *id.* at 13,841–42 (citing an internal study finding that “network-owned stations aired 4.3 hours more local news per week than did affiliates”). The FCC also cited an external study showing the differential to be 6.4 hours per week. *Id.* Rejecting the comments of the Project for Excellence in Journalism (PEJ), the FCC also concluded that network-owned stations produce higher-quality local news, as measured by industry awards. See *id.* at 13,842.

³²¹ *Id.* at 13,755. It also pointed out that “the average number of hours of local news and public affairs programming provided by the same-market cross-owned television-newspaper combinations was 25.6 hours per week, compared to 16.3 hours per week for the sample of television stations owned by a newspaper that is not in the same market as the station.” *Id.*

of more local news justified lifting the television/radio broadcast ownership ban.³²²

In sum, “local news” has become *the* critical component of the FCC’s “public interest” analysis, at least in the media ownership context. Although local news has long played an important role in the idea of “public service,”³²³ its predominance in the deregulation order is striking. The supervening norm that the FCC must pursue, the “public interest,” has now become practically identical to the number of hours of local news a station broadcasts.³²⁴ But what in fact is on the local news?

C. Local News

1. *Crime and Punishment.* — Violent crime. Crime occupies a heavy share of broadcast news programming.³²⁵ This is true for national news. It is also true for local news, which is “the most widely used source of information about crime.”³²⁶ The PEJ’s annual study of local news programming consistently finds that local newscasts spend about a quarter of their time on crime stories.³²⁷

Political scientists Frank Gilliam and Shanto Iyengar examined local news of a Los Angeles network affiliate for thirteen months, ran-

³²² See *id.* at 13,773–74 (attributing increased production to efficiencies).

³²³ See, e.g., Lorna Veraldi, *Newscasts as Property: Will Retransmission Consent Stimulate Production of More Local Television News?*, 46 FED. COMM. L.J. 469, 470 (1994) (pointing out that “[f]or half a century, lawmakers have searched for the carrot or the stick that would encourage American television stations to produce more local news and information programs,” including interpreting the “public interest” standard).

³²⁴ Beltway insiders and industry players may think me naïve in taking the Commission at its word. Local news may have been a convenient cover to justify certain regulatory changes sought for other reasons. Or it might have been a largely dissatisfying political compromise that captures no single Commissioner’s understanding of the “public interest.” Of course, the same could be said of the written opinions of appellate courts. But official explanations carry weight, not only for judicial opinions but also FCC reports. Departing from or ignoring these explanations and reasons in subsequent actions would be grounds for reversal under the “arbitrary and capricious” review of the Administrative Procedure Act. See, e.g., *Schurz Communications, Inc. v. FCC*, 982 F.2d 1043, 1053–54 (7th Cir. 1992) (finding new FinSyn rules arbitrary partly because the Commission did not explain its deviation from its prior 1983 tentative decision).

³²⁵ See ENTMAN & ROJECKI, *supra* note 261, at 80 (emphasizing the disproportionate broadcast of violent crimes, especially murder); Calvert, *supra* note 290, at 176–77; Barry C. Feld, *Race, Politics and Juvenile Justice: The Warren Court and the Conservative “Backlash”*, 87 MINN. L. REV. 1447, 1530 (2003) (“The mass media, especially the local electronic news media, overreport on the rarest types of crime, such as murder and rape: ‘If it bleeds, it leads.’” (quoting ENTMAN & ROJECKI, *supra* note 261, at 78)).

³²⁶ Franklin D. Gilliam, Jr. et al., *Crime in Black and White: The Violent, Scary World of Local News*, 1 HARV. INT’L J. PRESS/POL. 6, 7 (1996).

³²⁷ Tom Rosenstiel et al., *Local TV News: What Works, What Flops, and Why*, COLUM. JOURNALISM REV., Jan.–Feb. 1999, at 65, available at <http://archives.cjr.org/year/99/1/pej>. The PEJ is an affiliate of the Columbia University Graduate School of Journalism.

domly selecting a thirty-minute newscast, two days per week.³²⁸ The absolute number of minutes dedicated to crime stories was high. On average, there were three crime stories per day, accounting for 25% of the total minutes aired. In 51% of the newscasts, crime was the lead story.³²⁹ For comparison, the researchers obtained actual crime statistics during the corresponding time period. Although the ratio of violent crime arrests compared to all crime arrests in Los Angeles was 30%, violent crime news stories accounted for 78% of crime stories broadcast.³³⁰ Although the ratio of murder arrests compared to all arrests in Los Angeles was 2%, murder news stories accounted for 27% of all crime stories broadcast.³³¹

These figures do not in themselves demonstrate disproportionality because any such claim must provide a normative account of newsworthiness.³³² We should not presume that a one-to-one correspondence is “proportional.”³³³ For example, even if murders happen quite infrequently as compared to tax evasion, under what theory of newsworthiness should broadcasters have to provide thousands of minutes covering tax evasion for a single minute of murder? Still, the disparities suggest a lurid fascination with violent crime. They are also consistent with findings that the time allocated to crime stories does not correlate with changes in crime rates.³³⁴ These findings should not surprise us, given the strong financial incentives to focus on sensationalistic stories such as violent crimes. Financial success of broadcast stations requires high ratings, in order to sell more advertisements at

³²⁸ The station was KABC, affiliated with the ABC network. The time frame ranged from March 1993 to March 1994. Gilliam et al., *supra* note 326, at 8.

³²⁹ *Id.* at 8–9.

³³⁰ *Id.* at 9–10.

³³¹ *Id.* at 10; see also Feld, *supra* note 325, at 1530–31 (“Although homicide accounts for less than 1% of crime, it constitutes about one-quarter of all crime news coverage.”).

³³² Although, of course, many do make a disproportionality claim. See, e.g., LORI DORFMAN & VINCENT SCHIRALDI, OFF BALANCE: YOUTH, RACE & CRIME IN THE NEWS 4–5, 7 (2001), available at <http://www.buildingblocksforyouth.org/media/media.pdf> (“[D]epictions of crime in the news are not reflective of either the rate of crime generally, the proportion of crime which is violent, the proportion of crime committed by people of color, or the proportion of crime committed by youth.”).

³³³ It may be hard to imagine what a one-to-one correspondence between crime stories and all other stories even looks like. That requires the difficult and ambiguous task of identifying the universe of all possible news stories.

³³⁴ See Feld, *supra* note 325, at 1530 (“During the 1990s, overall violent crime decreased 20% while crime news coverage increased 83%, and homicides declined by one-third while network news coverage of them increased 473%.”); see also DORFMAN & SCHIRALDI, *supra* note 332, at 7 (after surveying over 100 studies, noting general consensus finding disparity between media depictions of crime and actual rates); *id.* at 10 (observing that while “there was a 1% increase in crime from 1979 to 1982, crime coverage in *Time* increased by 55% during that time period. Nationally, crime dropped by 20% from 1990 to 1998, while network TV showed an 83% increase in crime news. While homicide coverage was increasing on network news — 473% increase from 1990 to 1998 — homicides were down 32.9% from 1990 to 1998.” (citations omitted)).

higher rates. In turn, this strategy requires stations' pulling in more viewers any way they can.³³⁵

Violent crime news stories frequently involve racial minorities, especially African Americans.³³⁶ One reason is that racial minorities are arrested for violent crimes more frequently on a per capita basis than Whites.³³⁷ Given our social cognition review, we can predict what watching local news might do to us. If subliminal flashes of Black male faces can raise our frustration, as shown by the Computer Crash study, would it be surprising that consciously received messages couched in violent visual context have impact, too?³³⁸ In fact, we have already seen in the Mugshot study, described in the Introduction, that even ephemeral exposure to race can alter our opinions about crime and punishment.³³⁹ That study, also conducted by Gilliam and Iyengar, is one of the more sophisticated studies in a line of newscast experiments finding similar results.

³³⁵ See, e.g., Feld, *supra* note 325, at 1528 ("In an effort to increase audience shares and advertising revenues, local news programming favors an 'action news' format." (citing Gilliam et al., *supra* note 326, at 7)); *id.* at 1528 n.363 (citing ENTMAN & ROJECKI, *supra* note 261, at 44; and DARNELL M. HUNT, SCREENING THE LOS ANGELES "RIOTS": RACE, SEEING, AND RESISTANCE 31 (1997)) (emphasizing the need for advertising dollars). There may be other structural causes of the emphasis on violent crime in local news reporting, such as overreliance on police sources due to limited resources. According to Feld, the police may focus on violent crimes, and this focus may then be replicated in the media. See *id.* at 1534 (citing DORFMAN & SCHIRALDI, *supra* note 332, at 28). See generally LEONARD DOWNIE JR. & ROBERT G. KAISER, THE NEWS ABOUT THE NEWS: AMERICAN JOURNALISM IN PERIL 157-94 (2002) (describing what goes on in a local news room and explaining how overriding financial incentives encourage lower quality broadcasts, including sensationalist reports).

³³⁶ See DORFMAN & SCHIRALDI, *supra* note 332, at 13 (summarizing results from various studies).

³³⁷ See FBI, U.S. DEP'T OF JUSTICE, CRIME IN THE UNITED STATES 2002: UNIFORM CRIME REPORTS 17 (2003) (reporting the racial breakdown of violent crime arrestees: 59.7% White, 38.0% Black, 2.3% other), available at <http://www.fbi.gov/ucr/02cius.htm>. Whites constitute 75.1% of the population, while Blacks constitute 12.3% of the population. U.S. Census Bureau, *Race Alone or in Combination: 2000*, at http://factfinder.census.gov/servlet/QTTable?_bm=y&-qr_name=DEC_2000_SF1_U_QTP5&-lang=en (last visited Feb. 13, 2005). For murder arrestees during 2002, 47.7% were White and 50.0% were Black. FBI, *supra*, at 17. For similar statistics, see ENTMAN & ROJECKI, *supra* note 261, at 78-79; and Hurwitz & Peffley, *supra* note 141, at 376. This disproportionality in arrest rates is likely exacerbated, however, by unfavorable portrayals of Black criminals by local news media and concentration of news stories featuring blacks in violent crime stories. See *id.* (citing KATHLEEN HALL JAMIESON, DIRTY POLITICS: DECEPTION, DISTRACTION, AND DEMOCRACY (1992); and Robert M. Entman, *Blacks in the News: Television, Modern Racism and Cultural Change*, 69 JOURNALISM Q. 341 (1992)) ("These findings strongly suggest that the media contribution is one of both linking blacks to the issue of crime and, moreover, rendering stereotypes of blacks more negative.").

³³⁸ That impact can be on both explicit and implicit levels. For further evidence that conscious exposure to images can alter implicit bias, see *infra* pp. 1557-61.

³³⁹ See *supra* pp. 1491-92.

For example, using a similar experimental design,³⁴⁰ Gilliam had earlier found that exposure to a Black perpetrator mugshot produced statistically significant increases in two variables, defined as “Concern for Violent Crime”³⁴¹ and “Causal Attributions for Rising Crime.”³⁴² Another study has concluded that “judgments of the suspect’s guilt are significantly affected by the visual image of the race of the suspect, respondents’ stereotypes of blacks, and the interaction between these two variables.”³⁴³ A political scientist has even demonstrated how exposure to crime stories, depending on the race of the suspect, can alter self-reports of which presidential candidate participants will vote for.³⁴⁴

³⁴⁰ The researchers randomly exposed participants to a fifteen-minute videotaped local newscast that reported either a violent or a nonviolent crime. The report was digitally edited so that some of the participants saw a mugshot of a Black perpetrator, while others saw a White perpetrator. Gilliam et al., *supra* note 326, at 15. Afterward, the participants filled out a survey on crime attitudes that inquired about their level of fear of violent crime, their theories of why crime was rising, and how society might remedy the problem. *Id.* at 16.

³⁴¹ The corresponding survey question asked: “Lately, there has been a lot of attention to the problem of random street violence in Los Angeles. How serious of a problem do you think random street violence is in Los Angeles?” The response choices were “not at all serious,” “somewhat serious,” and “very serious.” *Id.* at 21 n.14.

³⁴² The corresponding survey question asked: “Please indicate whether you agree or disagree with the following explanations for the high rate of crime—‘breakup of the family’ and ‘declining influence of religion.’” The answer choices ranged from “strongly agree” to “strongly disagree.” *Id.*

³⁴³ Mark Peffley et al., *The Intersection of Race and Crime in Television News Stories: An Experimental Study*, 13 POL. COMM. 309, 315 (1996). In Peffley’s experiment, participants viewed a newscast of a violent crime. Although race was not mentioned in the dialogue, the perpetrator was *seen* being taken away in handcuffs. Some of the participants saw a White perpetrator, while others saw a Black perpetrator. *Id.* at 313. The participants were ninety-five White students in a political science class at a midwestern university. *Id.* One week earlier the participants had been instructed to complete a survey, which was an explicit (self-reported) measure of racial bias against African Americans. After viewing the video, the students answered various questions, such as assessing the likelihood of guilt. *Id.* at 313, app. at 327. Not surprisingly, those students who displayed the highest self-reported bias against Blacks were the most negative and punitive in their evaluations of the suspect in the Black perpetrator scenario. *Id.* at 321.

³⁴⁴ In one experiment, Nicholas Valentino showed participants a twelve-minute newscast that featured one crime story. Nicholas A. Valentino, *Crime News and the Priming of Racial Attitudes During Evaluations of the President*, 63 PUB. OPINION Q. 293, 301–02 (1999). The sample of 289 Los Angeles area adults was deemed “roughly comparable” to the area’s population in “gender, race, education, income, age, and partisanship.” *Id.* at 301. The news story shown to participants, which addressed gang activity, made no verbal mention of race. *Id.* at 301–02. However, a five-second mugshot of two suspects of the same race was edited in. The researchers showed some participants the mugshot, altered to reflect various racial groups, including White, Black, Asian, and Hispanic, and showed some participants no mugshot. They showed a control group no crime story at all. *Id.* at 302. After seeing this newscast, participants were asked to evaluate Bill Clinton and Bob Dole, who were the respective presidential nominees of the Democratic and Republican parties at the time of the study. *Id.*

In addition, participants answered the question: “In the upcoming presidential election in November, who are you likely to vote for?” *Id.* For the control group ($n = 31$), who saw no crime story, there was an 84% lead for Clinton over Dole. *Id.* at 305 fig.1. However, for those partici-

In the Mugshot study, Gilliam and Iyengar also used survey data to corroborate their experimental findings. In a large survey conducted at approximately the same time and location as the experiments, participants answered questions about their political opinions and media consumption habits.³⁴⁵ Three statistically significant correlations emerged: greater viewing of local news led to greater support for punitive remedies, more old-fashioned racism, and more “new racism.”³⁴⁶ Such results, confirmed in various contexts,³⁴⁷ should give us all pause. On the basis of this evidence alone, one could challenge the FCC’s unmindful adoration of local news as furthering the public interest — at least as local news is currently constituted.³⁴⁸ But the social science described in Part I suggests that far more might be at stake.

2. *Trojan Horse Viruses.* — I now make explicit what I have so far left implicit: local news programs, dense with images of racial minorities committing violent crimes in one’s own community,³⁴⁹ can be

pants who saw the crime story and saw either a White suspect or no suspect at all ($n = 118$), Clinton’s lead fell to 60%. *Id.* Finally, for those who saw a minority perpetrator ($n = 137$), Clinton’s lead was reduced to 40%. *Id.* All these results were statistically significant at $p < 0.01$. *Id.* at 306. These results are impressive, in part because nothing in the newscast directly mentioned the presidential race. *See id.* at 313.

³⁴⁵ *See* Gilliam & Iyengar, *supra* note 6, at 570.

³⁴⁶ *See id.* at 571 tbl.5 (reporting that at the $p < 0.01$ level, there was a 4% increase for punitive remedies, a 4% increase in old-fashioned racism, and a 7% increase in “new racism”). Gilliam and Iyengar characterize “new racism” as “symbolic, subtle, covert, hidden, or underground.” *Id.* at 566 (internal quotation marks omitted); *see also* James M. Avery & Mark Peffley, *Race Matters: The Impact of News Coverage of Welfare Reform on Public Opinion*, in RACE AND THE POLITICS OF WELFARE REFORM 131, 136 (Sanford F. Schram et al. eds., 2003) (internal citations omitted) (“Experimental evidence suggests that even a brief visual image of a black male in a typical nightly news story on crime is powerful and familiar enough to activate viewers’ negative stereotypes of blacks, producing racially biased evaluations of black criminal suspects. In their experimental studies manipulating the skin color of a male perpetrator in a local news broadcast, Gilliam and associates found that when the perpetrator was African American, more subjects endorsed punitive crime policies and negative racial attitudes after watching the news broadcast.”).

³⁴⁷ In addition to crime, the link between race and welfare reform has been extensively studied by political scientists. There is convincing data that people respond differently to news reports about welfare based on whether a Black or White person is depicted as the welfare recipient. *See, e.g.,* Avery & Peffley, *supra* note 346, at 140–42 (finding that when welfare reform was described as a failure, whether the picture depicted a Black woman or White woman affected how much she was blamed for losing her job).

³⁴⁸ To be clear, my critique is with the way that local news is currently constituted, with its surfeit of violent crime stories. No doubt watching local news, even as it exists today, produces substantial benefits to the citizenry. Accordingly, I am calling for a fuller accounting that considers not only the benefits of local news but also the costs.

³⁴⁹ The use of visual images allows racial mapping in ways that newspapers often do not. *See* ENTMAN & ROJECKI, *supra* note 261, at 89. This characteristic of visual images raises important questions about how and when newspapers decide to run photographs of suspects, identify race in textual descriptions, or publish names (which would also allow racial mapping). *See generally* GANDY, *supra* note 264, at 159–64 (discussing blindspots, stigmas, and stereotypes of mass media in race matters).

analogized to Trojan Horse viruses. A type of computer virus, a Trojan Horse installs itself on a user's computer without her awareness. That small program then runs in the background, without the user's knowledge, and silently waits to take action — whether by corrupting files, e-mailing pornographic spam, or launching a “denial of service” attack — which the user, if conscious of it, would disavow.

Typically, a Trojan Horse comes attached secretly to a program or information we actively seek. For instance, we might download a new program for a trial run, and embedded inside may be a Trojan Horse that installs itself without our knowledge. Or, we might browse some website in search of information, and a small javascript bug may be embedded in the page we view. Here is the translation to the news context: we turn on the television in search of local news, and with that information comes a Trojan Horse that alters our racial schemas. The images we see are more powerful than mere words.³⁵⁰ As *local* news, they speak of threats nearby, not in some abstract, distant land.³⁵¹ The stories are not fiction but a brutal reality.³⁵² They come from the most popular and trusted source.³⁵³

³⁵⁰ See, e.g., BARRIE GUNTER, POOR RECEPTION: MISUNDERSTANDING AND FORGETTING BROADCAST NEWS 71–72 (1987) (reporting studies demonstrating that visuals are better recalled than words and may dominate the verbal content of television news); Avery & Peffley, *supra* note 346, at 135 (collecting sources suggesting that visual images “trump auditory and text messages”).

³⁵¹ See Robert M. Entman, *Modern Racism and the Images of Blacks in Local Television News*, 7 CRITICAL STUD. MASS COMM. 332, 335 (1990) (identifying racial messages delivered through local news as “anchored close to home and in all likelihood help[ing] to construct the audience’s sense of community well-being and community threat in a way that the more distant and abstracted national news cannot”).

³⁵² Cf. ENTMAN & ROJECKI, *supra* note 261, at 78 (arguing that nonfictional violence is often more graphic than the violence shown on broadcast entertainment programs); Eunkyung Jo & Leonard Berkowitz, *A Priming Effect Analysis of Media Influences: An Update*, in MEDIA EFFECTS: ADVANCES IN THEORY AND RESEARCH 43, 54–55 (Jennings Bryant & Dolf Zillman eds., 1994) (describing a study in which children “who watched a realistic news portrayal of a fight scored higher on an aggression index than other youngsters who had seen the same story in a fantasy entertainment condition or than the control group shown an ordinary product commercial”). Some have also suggested a causal relationship between local news and entertainment programming: the more violent the former, the more violent the latter. See John M. Higgins, *Blame It on Local News*, BROADCASTING & CABLE, Dec. 9, 2002, at 3 (“HBO President Chris Albrecht and other programming executives justify the increasing amount of graphic violence they’re putting on their networks [in this way]: Local news stations are numbing audiences to violence; hence, they’re more accepting of violence in their dramas.”).

³⁵³ See, e.g., ENTMAN & ROJECKI, *supra* note 261, at 79 (citing PEW RESEARCH CTR. FOR THE PEOPLE AND THE PRESS, FEWER FAVOR MEDIA SCRUTINY OF POLITICAL LEADERS: PRESS “UNFAIR, INACCURATE AND PUSHY”, <http://people-press.org/reports/print.php3?ReportID=112> (Mar. 21, 1997); and Dan Trigoboff, *Survey: Viewers Applaud Their Local News*, BROADCASTING & CABLE, Dec. 21, 1998, at 60). Entman and Rojecki report that 72% “of Americans watch local news regularly,” as compared to 41% for network news. *Id.* (quoting PEW RESEARCH CTR., *supra*) (internal quotation marks omitted) (citing Trigoboff, *supra*). Also, local news gets higher ratings than network news in many markets. *Id.*; see also Entman, *supra* note

Two clarifications are warranted. First, in the computer world, Trojan Horse viruses are written by programmers and unleashed into the wild to do the authors' bidding. That is not my account of race and racism. For example, I do not believe that broadcast licensees or local news producers are purposefully designing visceral images of minority-perpetrated crimes to implant viruses in the audience's brains that harm racial minorities.³⁵⁴ One could plausibly accuse political advertisers of doing precisely this, when they launch Willie Horton-like ads³⁵⁵ or their more sophisticated equivalents.³⁵⁶ But no such case can be made against local news producers generally, who are merely trying, on a charitable view, to provide satisfactory news while maintaining top ratings in increasingly competitive media environments.³⁵⁷

Second, viruses are executable code, programs that take inputs and generate outputs. Isn't news, or any other media input, something different — just data? Doesn't my account blur the data/program distinction? Yes, it does, but that distinction may not be tenable if our minds function partly like neural networks, which they seem to do. As suggested earlier, my racial schemas model can be seen as a higher-level metaphor for a lower-level description that models the brain as a neural network, a massively distributed mesh of nodes with varied

351, at 335 (discussing the greater popularity of local news in the Chicago market); Gilliam & Iyengar, *supra* note 6, at 560 (noting that local news can be seen "almost anytime — mornings, afternoons, evenings, prime time, and late night" and suggesting that local news' dominance has been variously demonstrated). But local news ratings are also headed downward. See DOWNIE & KAISER, *supra* note 335, at 172 (noting that a survey from the Pew Research Center found that "only 56 percent of Americans were watching local television news regularly in 2000, down from 77 percent in 1993").

As compared to network news, local news is evaluated as better in quality and more credible by viewers. See Calvert, *supra* note 290, at 171 (citing Lawrence K. Grossman, *Why Local TV News Is So Awful*, COLUM. JOURNALISM REV., Nov.-Dec. 1997, at 21).

³⁵⁴ In fact, news reporters are more likely to identify themselves as politically liberal than is the average American. See S. ROBERT LICHTER ET AL., *THE MEDIA ELITE* 21 (1986) (providing results of interviews with 238 journalists at prestigious news publications in 1979 and 1980 that showed 54% identified as politically liberal versus 17% as politically conservative); see also *infra* note 405.

³⁵⁵ Tali Mendelberg has carefully demonstrated that exposure to this now infamous advertisement during the 1988 presidential campaign implicitly triggered racial resentment (what I would call "negative racial meanings"), which in turn influenced the election. See TALI MENDELBERG, *THE RACE CARD: CAMPAIGN STRATEGY, IMPLICIT MESSAGES, AND THE NORM OF EQUALITY* 134-68 (2001) (describing the various phases of the Willie Horton campaign); *id.* at 173-83 (demonstrating empirically the advertisements' impact on public sentiment).

³⁵⁶ See generally Nicholas A. Valentino et al., *Cues That Matter: How Political Ads Prime Racial Attitudes During Campaigns*, 96 AM. POL. SCI. REV. 75 (2002) (demonstrating that political advertisements that make no verbal mention of race but that use racial images can activate racial attitudes that even influence viewers' opinions on subjects not especially connected to race, such as the federal budget and taxes).

³⁵⁷ Then again, how should we characterize photo editors at *Time* who purposefully darkened O.J. Simpson's skin color during his murder trial? See Howard Kurtz, *Time's 'Sinister' Simpson: Cover Photo Was Computer-Enhanced*, WASH. POST, June 22, 1994, at D1.

strengths in linkages among them.³⁵⁸ Neural nets are not programmed by writing a linear, flowchart-like order of instructions for execution. Rather, they are trained through exposure to countless cases.³⁵⁹ Through some learning rule, such as back propagation, those connections that produce a correct answer are subtly strengthened, and those that produce a wrong answer are subtly weakened.³⁶⁰ After sufficient numbers of exposures, the neural network is deemed “trained,” and its knowledge/programming is reflected in the relative strengths of the links among the various nodes. When the brain is modeled this way, much of the data/program distinction dissolves: a part of us is in fact programmed by what we see.

How do we know violent crime stories can, like Trojan Horses, exacerbate implicit bias? The Mugshot study and other work by political scientists using the newscast paradigm are suggestive. Further evidence comes from studies that demonstrate media primings of racial schemas. For example, we now know that exposure to violent rap music can increase implicit bias against African Americans³⁶¹ and that

³⁵⁸ See *supra* pp. 1498–99 & note 41.

³⁵⁹ Cf. Lawrence, *supra* note 28, at 323 (describing how children learn racism through unarticulated “tacit understandings”).

³⁶⁰ See RAYMOND KURZWEIL, *THE AGE OF SPIRITUAL MACHINES* 77 (1999).

³⁶¹ Laurie Rudman and Matthew Lee found that a thirteen-minute audio-only exposure to violent, misogynistic rap music increased the implicit racial bias of participants, as measured by the IAT. See Laurie A. Rudman & Matthew R. Lee, *Implicit and Explicit Consequences of Exposure to Violent and Misogynous Rap Music*, 5 *GROUP PROCESSES & INTERGROUP REL.* 133, 137–38 & tbl.1 (2002). Those participants in the “prime” condition listened to rap music that portrayed African Americans as violent and sexist. *Id.* at 136 (providing lyric samples). Those in the “control” group listened to contemporary pop tunes. *Id.* Then, the participants took a stereotype IAT, categorizing Black versus White names (for example, Jamal versus Hank) and negative versus positive words (for example, hostile versus calm). See *id.* at 136–37. Primed subjects generated higher IAT scores. See *id.* at 137 (reporting scores of $M = +327$ ms for primed subjects versus $M = +107$ ms for control subjects at the $p < 0.001$ level).

Not surprisingly, the rap music prime also displayed increased measures of *explicit* bias on a self-reported stereotype endorsement scale. *Id.* at 138 tbl.1. However, the only statistically significant increase was in participants identified as “high prejudice” according to the MRS. See *id.* This evidence of priming and dissociation indicates that “low prejudice” people can sincerely claim that rap music does not influence their explicit agreement with racial stereotypes; nevertheless, like a Trojan Horse, the audio input will at least temporarily increase their implicit bias. See *id.* at 145 (“[E]ven low prejudiced people are unlikely to recognize the power of the situation and implicit stereotypes when they make interpretative judgments about others.” (citation omitted)); see also James D. Johnson et al., *Converging Interracial Consequences of Exposure to Violent Rap Music on Stereotypical Attributions of Blacks*, 36 *J. EXPERIMENTAL SOC. PSYCHOL.* 233, 238 (2000).

In the study by Johnson et al., researchers exposed participants to a four minute audio-only rap song with extremely violent lyrics and sound effects, a four minute audio-only nonviolent rap song, or no music at all. Both songs were attributed to Black musicians. *Id.* at 237, 239. On subsequent interpretive tasks, participants who heard the violent music evaluated Black targets in schema-consistent ways: Those who listened to the violent music were more likely than those in the other two groups to attribute Black male violence to disposition rather than situation. They

playing the video game *Doom* can increase one's implicit self-concept of aggressiveness³⁶² — all the while having no statistically significant impact on one's explicit, self-reported views. Still further evidence comes indirectly from research Nilanjana Dasgupta calls the “third wave” of implicit bias research, which examines the malleability of implicit bias.³⁶³ This research demonstrates that implicit bias can be exacerbated or mitigated by the information environments we inhabit.

Positive Role Models. Consider, for example, how exposure to positive exemplars of subordinated categories can decrease implicit bias. Nilanjana Dasgupta and Anthony Greenwald found that implicit attitudes could be changed without conscious effort simply by exposing people to particular types of content.³⁶⁴ Participants were first given a “general knowledge” questionnaire.³⁶⁵ For the pro-Black condition group, the researchers used names and images of positive Black exemplars, such as Martin Luther King, Jr., and negative white exemplars, such as Jeffrey Dahmer.³⁶⁶ For the pro-White condition group, the valences of the images were reversed (Louis Farrakhan and

were also more likely to view Blacks as less qualified for a job requiring intelligence. *See id.* at 242–44 & tbl.1. These effects were seen in both Black and White participants. *See id.*

³⁶² Eric Uhlmann and Jane Swanson examined how playing violent video games might alter one's self-concept of aggressiveness. *See generally* Eric Uhlmann & Jane Swanson, *Exposure to Violent Video Games Increases Automatic Aggressiveness*, 27 J. ADOLESCENCE 41 (2004). One hundred and twenty-one psychology students participated in the experiment, in which one group played the violent first-person shooter video game *Doom* for ten minutes and another group played *Mahjongg: Clicks*, an absorbing puzzle game. *Id.* at 43. Participants then took an IAT measuring the implicit connections between Self and Other and Aggressive and Peaceful. *Id.* at 44. Finally, participants answered explicit self-reports about their own aggressiveness. *Id.* Participants who played *Doom* implicitly associated themselves more with the concept of aggressiveness than did those who played *Mahjongg: Clicks*. *See id.* at 46 (reporting that at the $p = 0.036$ level, “participants in the *Doom* condition were more likely to automatically associate themselves with aggression ($M = -130$ ms, s.d. = 153 ms) than participants in the *Mahjongg* condition ($M = -201$ ms, s.d. = 204 ms), a difference that was statistically significant”). Although women had implicit self-concepts that were more peaceful than men ($p = 0.023$), there was no interaction between participant gender and game condition. *See id.* Moreover, *Doom* had no impact on explicit self-reports. *See id.* at 47. Collectively, this and the music study discussed above, *see supra* note 361, suggest that the electronic media we encounter can activate certain schemas and at least temporarily increase certain implicit associations.

³⁶³ *See* Dasgupta, *supra* note 100, at 160–62.

³⁶⁴ Nilanjana Dasgupta & Anthony G. Greenwald, *On the Malleability of Automatic Attitudes: Combating Automatic Prejudice With Images of Admired and Disliked Individuals*, 81 J. PERSONALITY & SOC. PSYCH. 800, 807 (2001).

³⁶⁵ Of the 48 participants, 31 were White and 17 were Asian. *Id.* at 802.

³⁶⁶ The complete list of positive Black images included Martin Luther King, Jr., Jesse Jackson, Colin Powell, Denzel Washington, Eddie Murphy, Michael Jordan, Tiger Woods, Will Smith, Bill Cosby, and Gregory Hines. For negative White images, Dasgupta and Greenwald used Ted Bundy, Jeffrey Dahmer, Timothy McVeigh, Charles Manson, Al Capone, Ted Kaczynski, Terry Nichols, Howard Stern, John Gotti, and John Dillinger. *See id.*, app. B at 811. Obviously, the choice of these images could be debated.

John F. Kennedy, for example).³⁶⁷ Finally, for a control group, the questionnaire required correct identification of insects and flowers. After finishing the questionnaire, participants took an IAT³⁶⁸ and then completed a survey of racial bias.³⁶⁹

The type of questionnaire had no impact on participants' explicit bias as measured by the self-reports. By contrast, the researchers found that the questionnaires had a surprisingly significant effect on implicit bias as measured by the IAT: those participants who had experienced the pro-Black condition reduced their implicit bias by more than half.³⁷⁰ These results persisted for over twenty-four hours, as measured by a follow-up test.³⁷¹

The authors explained the results in terms of exemplar accessibility.³⁷² When we evaluate social groups, we do so by calling on particular exemplars of that group retrieved from memory. The cache of racial meanings associated with a racial category may have internally inconsistent content, which may be maintained through the use of sub-categories, subtypes, or different exemplars; which exemplar gets activated depends partly on which exemplar is most accessible.³⁷³ Thus, recent priming through visual images (such as the pictures presented in the "general knowledge" questionnaire) can alter the accessibility of one exemplar over another (Martin Luther King, Jr., over Louis Farrakhan, Michael Jordan over Mike Tyson).³⁷⁴ The control group, which answered a questionnaire about flowers and insects, produced results that were indistinguishable from the pro-White condition, which sug-

³⁶⁷ The negative Black images included O.J. Simpson, Mike Tyson, Louis Farrakhan, Marion Barry, Arthur Washington, Lonny Gray, Tyshawn Williams, Charles Brackett, Michael McClinton, and Stanley Obas. Positive White images included Clint Eastwood, Jim Carrey, Tom Cruise, David Duchovny, Tom Hanks, Jay Leno, John F. Kennedy, Robert Redford, Norman Schwarzkopf, and Peter Jennings. *See id.*, app. B, at 812.

³⁶⁸ The IAT used stereotypically White or Black names, not images. *See id.* at 803.

³⁶⁹ These surveys included a feeling thermometer as well as semantic differential scales. *Id.*

³⁷⁰ The net decrease in latency came from faster reaction times for the "Black + pleasant" and the "White + unpleasant" combinations in the IAT. Interestingly, the latencies for the "White + pleasant" and the "Black + unpleasant" combinations did not change across the various exemplar conditions. *See id.* at 807.

³⁷¹ *See id.*

³⁷² *See id.* at 807–08 ("Because the exemplars selected for our studies were famous and infamous individuals who were known to participants beforehand, it is unlikely that exposure to these exemplars over the course of the experiment elicited the creation of new mental representations. More likely, exposure simply increased the accessibility of known exemplars."). In this article, I have steadfastly tried to avoid the controversies in social cognition among those who favor prototype, exemplar, classical schema theory, or neural network models. My racial mechanics model, which is articulated at a high level of abstraction, is perfectly consistent with this exemplar accessibility story.

³⁷³ *See id.*

³⁷⁴ *See id.*

gests that the default “exemplar” or racial meaning for Whites is favorable.³⁷⁵

The dissociation between explicit and implicit bias deserves highlighting. Again, the researchers found discrepancy between explicit and implicit attitude measures. For the explicit surveys, participants could discount the set of images provided as atypically favorable or unfavorable and consciously correct their attitudes before answering.³⁷⁶ By contrast, the more automatic processes, as measured by the IAT, seemed not to be subject to conscious compensation or discounting.³⁷⁷ Such dissociation recommends a dual-pronged approach to addressing problematic racial schemas: those strategies that correct for explicit bias may not be particularly useful as a solution to implicit bias, and vice versa.³⁷⁸

Mental Imagery. A study by Irene Blair, Jennifer Ma, and Alison Lenton focusing on counterstereotypic mental imagery is also telling. Motivated by evidence that visualization shares many characteristics with real experiences and thus can influence learning and behavior, they tested whether mental imagery could moderate implicit stereotypes.³⁷⁹ Individuals instructed to visualize a counterstereotypic image would, in effect, be priming themselves in a way that would make counterstereotypic actions easier.

In the first experiment,³⁸⁰ one group of participants was instructed to spend a few minutes imagining a strong woman, her attributes and abilities, and the hobbies she enjoys; another group was asked to imagine a Caribbean vacation.³⁸¹ Those who imagined the strong woman registered a significantly lower level of implicit stereotype in the IAT.³⁸² In subsequent experiments, additional groups were added for

³⁷⁵ See *id.* at 805.

³⁷⁶ See *id.* at 808.

³⁷⁷ See *id.*

³⁷⁸ See *id.* (suggesting that explicit attitudes might best be addressed through “deep cognitive processing,” whereas implicit bias may benefit from frequent application of “shallower processing”). In a follow-up study, Dasgupta made similar findings on a gender stereotype IAT. See Nilanjana Dasgupta & Shaki Asgari, *Seeing Is Believing: Exposure to Counterstereotypic Women Leaders and Its Effect on the Malleability of Automatic Gender Stereotyping*, 40 J. EXPERIMENTAL SOC. PSYCHOL. 642, 647 (2004) (finding that women who had been primed with “famous female leaders were significantly faster at associating women with leadership attributes (Mean RT = 772 ms) compared with those who had previously seen control exemplars [flowers] (Mean RT = 860 ms)”).

³⁷⁹ Irene V. Blair et al., *Imagining Stereotypes Away: The Moderation of Implicit Stereotypes Through Mental Imagery*, 81 J. PERSONALITY & SOC. PSYCHOL. 828, 828–29 (2001).

³⁸⁰ Researchers selected forty-two undergraduates — seventeen male and twenty-five female — as participants. See *id.* at 830.

³⁸¹ See *id.*

³⁸² See *id.* at 831. For the neutral imagery group, the reaction time difference between the schema-consistent and schema-inconsistent blocks was ninety-five milliseconds. For the counter-

comparison, including a group asked to imagine stereotypic “feminine women, such as storybook princesses or Victorian women,”³⁸³ and a group that engaged in no imagery whatsoever.³⁸⁴ There were statistically significant differences in the IATs among the counterstereotypic,³⁸⁵ no imagery/neutral imagery, and stereotypic³⁸⁶ groups, with reaction time differences increasing in that order.³⁸⁷ In further experiments, Blair and colleagues employed different measures of implicit stereotypes, including the more recently invented Go/No-go Association Test (GNAT)³⁸⁸ and a false memory measure.³⁸⁹ Based on the results of these experiments,³⁹⁰ the researchers concluded that there is “little doubt that the [counterstereotypic] mental imagery per se was responsible for diminishing implicit stereotypes.”³⁹¹

stereotypic imagery group, the difference was twenty-four milliseconds, which reached statistical significance at $p < 0.05$. *See id.* at 831 tbl.1.

³⁸³ *See id.* at 833 (internal quotation marks omitted).

³⁸⁴ The researchers found no statistically significant difference between the no-imagery and the neutral-imagery (Caribbean vacation) groups. *See id.* at 832.

³⁸⁵ *See id.* (showing difference between counterstereotypic group and neutral/no imagery group at $p < 0.05$).

³⁸⁶ *See id.* at 833 (showing difference between neutral and stereotypic group at $p < 0.05$).

³⁸⁷ What is specifically imagined is important. In one experiment, participants in the stereotypic imagery group were asked to imagine a “weak woman” characterized as a “somewhat fragile [woman that] might be described as a ‘delicate flower.’” *Id.* at 831. Surprisingly, this group showed differences in IAT scores that were statistically insignificant from the scores of the no-imagery group. *Id.* at 832. The researchers hypothesized that this result was due to the imaging instructions, which encouraged the mental image of an insecure, dependent woman. This image did not match the specific stereotypic words used in the IAT, which included “delicate,” “fine,” and “dainty.” *See id.* By switching the imaging directions to the “storybook princess,” researchers saw the increase in IAT that they had predicted. *See id.* at 833.

³⁸⁸ The GNAT measures an attitude or stereotype for a single category. By contrast, the IAT requires pairs of categories, such as female/male (as mapped to supportive/agentive). This use of pairs sometimes raises questions regarding whether the reaction time differential is produced by the association that female is more supportive or less agentive. The GNAT avoids such confounds. *See id.* at 834; *see also* Devos & Banaji, *supra* note 103, at 15–16 (describing differences between the GNAT and IAT).

³⁸⁹ The false memory test measured the rate at which participants incorrectly “remembered” being presented with words that are stereotypically associated with those that were actually presented. *See* Blair et al., *supra* note 379, at 835–36.

³⁹⁰ These last two experiments revealed several intriguing results. For instance, in the GNAT tests, men showed no difference regardless of their imaging. Women, by contrast, did show a statistically significant difference between the “strong woman” condition and others. *See id.* at 835 tbl.4. This result was obtained notwithstanding the fact that men had no difficulty producing counterstereotypic images of women, as reflected by independent review of the described mental images they generated. *See id.* at 835.

³⁹¹ *Id.* at 837.

These studies suggest that the images consumed matter.³⁹² Specifically, consuming positive images can decrease individuals' implicit bias, although they may register no difference on measures of explicit bias. Conversely, it seems reasonable to suppose that consuming negative images can exacerbate implicit bias. Recall the group in the Blair study instructed to imagine stereotypic women. And if mental imagery can produce such effects, watching direct portrayals in electronic media may well have an even stronger impact.

Coed Education. For those who are rightly skeptical about external validity — translating laboratory findings into real-world results — there is now some evidence that exposure to counterstereotypic exemplars decreases implicit bias in real-world situations. Nilanjana Dasgupta and Shaki Asgari performed a longitudinal study of female students before and after their first year of college.³⁹³ Half the participants were recruited from a coeducational college, whereas the other half attended a women's college. Both groups took tests measuring explicit and implicit bias and completed campus experience questionnaires.³⁹⁴ The two groups started with statistically indistinguishable levels of *implicit* bias: both groups viewed women stereotypically, as more "supportive" than "agentic."³⁹⁵ What happened after one year of college? On average, the implicit bias of those who had attended women's colleges disappeared. By contrast, the implicit bias of those who had attended coeducational colleges increased.³⁹⁶ Providing further evidence of dissociation, the groups' *explicit* self-reported endorsements of stereotypes did not change regardless of the college attended or time of measurement.³⁹⁷

³⁹² See Paul G. Davies et al., *Consuming Images: How Television Commercials that Elicit Stereotype Threat Can Restrain Women Academically and Professionally*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 1615 (2002). In this stereotype threat study, male and female participants were exposed to a set of real TV commercials that was either stereotypic or counterstereotypic. Both sets featured four identical commercials deemed neutral, without any human beings. The stereotypic set featured two additional commercials, with a woman excited about acne medicine in one case and brownies in the other. The counterstereotypic set featured two different commercials, with a woman depicted as an expert about automobiles in one case and healthcare in the other. *Id.* at 1619. After being primed with these commercials, the participants took a lexical-decision task to measure stereotype activation, then a difficult GRE-based math test scored the standard way. *Id.* Women and men scored equally well when primed with the counterstereotypic ads ($M = 31\%$ for women; $M = 34\%$ for men). In sharp contrast, in the stereotypic condition, women performed far worse than the men ($M = 19\%$ and $M = 39\%$ respectively; $p < 0.01$). *Id.* at 1620.

³⁹³ See generally Dasgupta & Asgari, *supra* note 378.

³⁹⁴ See *id.* at 649–50.

³⁹⁵ See *id.* at 651.

³⁹⁶ The IAT effect for those attending a women's college started at 31 ms and went down to -5 ms. By contrast, the IAT effect for those attending a coed college started at 74 ms and went up to 128 ms. See *id.*

³⁹⁷ See *id.* at 653.

But what was the mediating variable? On the basis of the campus experience questionnaire, the researchers ran regressions to see which, if any, campus environmental factors correlated with the change in implicit bias.³⁹⁸ Of the variables measured, the only statistically significant correlation was to “exposure of female faculty” (and not, for example, number of courses taken with gender-related content, say in the women’s studies department).³⁹⁹ Although this longitudinal field study explored the social schema of gender, we should not be surprised to find similar results soon on race. In the end, although we may not be able to directly command ourselves to show less implicit bias on the IAT,⁴⁰⁰ we may be able to do so indirectly by altering our informational and interactional environment.⁴⁰¹ This solves the riddle of how automatic processes may resist direct conscious control but nevertheless be mutable.⁴⁰²

To summarize: Local news provides data that we use consciously in a rational analysis to produce informed opinions on, say, criminal pun-

³⁹⁸ The questionnaire results were converted into eight variables: percentage of course instructors who were women, percentage of courses taken in which male students typically outnumber female students, percentage of courses taken with gender-related content, supportiveness of faculty and staff, faculty or staff viewed as role models, female role models on campus, degree of belief that college was preparing them for future leadership roles, and gender-issue-related extracurricular activities. *See id.* at 650.

³⁹⁹ *Id.* at 651 ($p = 0.004$). Initially, the number of math and science courses taken also seemed to produce a significant effect. However, the researchers determined that this was caused by the fact that math and science courses at coeducational colleges were disproportionately taught by male faculty. Controlling for the effect of number of female faculty, the interaction between math-science courses and implicit bias lost significance. *See id.* at 652–53.

⁴⁰⁰ *But cf.* Armour, *supra* note 50, at 768–69 (suggesting explicit instructions to juries to counter implicit biases). Armour describes prejudice as similar to a “bad habit,” which must be identified and consciously resisted. *Id.* at 756–57. On the one hand, there is strong evidence that an individual’s motivation, for example to be impartial, can influence how schemas parse information. *See supra* note 43 (discussing the “motivated tactician” model). On the other hand, there is disheartening information about ironic processes, in which attempts to suppress biases fail or backfire. *See supra* note 69. To the extent the legal literature has already called for explicit cognitive processes to engage with bias at all levels, the contribution of my article is to emphasize how implicit biases could be engaged on implicit levels. This strategy fights implicit fire with implicit fire.

⁴⁰¹ *Cf.* Aberson, *supra* note 194, at 345 (finding correlations between outgroup friendships and lower implicit bias).

⁴⁰² *See* Banaji, *supra* note 108, at 138–39 (noting the paradox that “[c]onscious and unconscious forms of evaluation are both independent and associated”); *see also* Jason P. Mitchell et al., *Contextual Variations in Implicit Evaluation*, 132 J. EXPERIMENTAL PSYCHOL.: GENERAL 455 (2003). In Mitchell’s study, the researchers had each participant identify popular Black athletes and unpopular White politicians. IATs were run using images from these two categories. When the participants were asked to focus on profession, they predictably showed substantial bias in favor of Black athletes. *Id.* at 459. By contrast — and this is what is surprising — when the participants were asked to focus on race, they showed a mild preference in favor of the White politicians notwithstanding the fact that they were chosen because they were unlikable. *Id.* According to the authors, this was the “first such study to demonstrate a sharp, rapid attitude dissociation as a function of the attended category.” *Id.* at 461.

ishment. But these newscasts also activate and strengthen linkages among certain racial categories, violent crime, and the fear and loathing such crime invokes. In this sense, the local news functions precisely like a Trojan Horse virus. We invite it into our homes, our dens, in through the gates of our minds, and accept it at face value, as an accurate representation of newsworthy events. But something lurks within those newscasts that programs our racial schemas in ways we cannot notice but can, through scientific measurements, detect. And the viruses they harbor deliver a payload with consequences, affecting how we vote for “three strikes and you’re out” laws, how awkwardly we interact with folks, and even how quickly we pull the trigger.

3. *The Accuracy Objection.* — A predictable objection is that the violent content, including crime committed by racial minorities, is a feature, not a bug. In other words, the data presented are not skewed and instead faithfully reflect a reality that the local news did not create. I have three responses to this “accuracy objection”: the data are likely not fairly presented; our memories and abilities to see patterns are selective; and we interpret the data in self-serving ways.

First, the information broadcast is probably not fair and balanced. There is a *prima facie* case that the local media give disproportionate attention to violent crime, in which Black suspects feature prominently.⁴⁰³ Furthermore, ample evidence shows that the media treats Black-perpetrator stories differently, representing and portraying suspects in a more threatening manner than comparable White perpetrators.⁴⁰⁴ Specifically, Robert Entman explains that because of production biases in local newscasts, Black suspects are more likely to remain unnamed and in physical custody, and less likely to speak for themselves.⁴⁰⁵ As a result, while there is evidence that the statistical

⁴⁰³ See *supra* pp. 1549–51; see also Feld, *supra* note 325, at 1533 (suggesting that the proximity of news stations to urban, largely minority population centers makes it easier to produce violent crime stories).

⁴⁰⁴ See, e.g., JAMIESON, *supra* note 337, at 133–34 (concluding that national news from 1985 to 1989 was more likely to show Blacks restrained and in mugshots, as compared to Whites); Ted Chiricos & Sarah Eschholz, *The Racial and Ethnic Typification of Crime and the Criminal Typification of Race and Ethnicity in Local Television News*, 39 J. RES. IN CRIME & DELINQ. 400, 400 (2002) (reviewing local news from three Orlando, Florida, television stations and concluding that although the media did not quantitatively overrepresent minorities, they did qualitatively frame stories such that Blacks and Hispanics appeared more threatening than similarly situated Whites); Gilliam et al., *supra* note 326, at 13–14 & fig.5 (identifying differences between coverage of White and Black perpetrators, in terms of the use of “hostile sources” and “hostile newscaster comments”); cf. ENTMAN & ROJECKI, *supra* note 261, at 81 (suggesting that White victim stories are given more time, perhaps due to class bias); Feld, *supra* note 325, at 1536 (suggesting that White victimization is considered more newsworthy).

⁴⁰⁵ See Robert M. Entman, *Blacks in the News: Television, Modern Racism and Cultural Change*, 69 JOURNALISM Q. 341, 349–53 (1992); see also Feld, *supra* note 325, at 1525 (suggesting that “[n]ews professionals may unconsciously cater to the stereotypes of their audience” (emphasis omitted)). Although I have focused on racial meanings of violence and criminality, there are other

prominence of Blacks portrayed in crime news is “not that much out of line with the actual Black arrest rate,”⁴⁰⁶ the emphasis on violent crime appears to skew public perceptions.⁴⁰⁷

Second, even if local news accurately reflected reality, we see “illusory correlations.”⁴⁰⁸ Whenever two salient events are noticed together, that combination leaves a deep impression in our memories and leads us to overestimate its frequency.⁴⁰⁹ Because racial minorities are numerical minorities (and therefore often salient) and because bad acts (for example, crimes) are also unusual and salient, when racial minorities commit bad acts, the information gets more deeply imprinted and weighted than is statistically warranted.⁴¹⁰ In other words, even if the exact same percentage of Whites and Blacks commit a bad act, and the data are provided objectively, we will (mis)remember Blacks committing a disproportionately higher number of bad acts. These findings have been confirmed in numerous experiments.⁴¹¹ This bias can couple with even more dramatic memory errors. For instance, Gilliam and Iyengar have demonstrated that even when a crime news story does not show any perpetrator and provides no information about

meanings as well, such as poverty and welfare. Martin Gilens has done substantial work in this field, demonstrating that the news portrays poverty as a Black problem. See MARTIN GILENS, *WHY AMERICANS HATE WELFARE: RACE, MEDIA, AND THE POLITICS OF ANTIPOVERTY POLICY* 149 (1999) (“Despite the liberal political attitudes of news professionals and the liberal racial views expressed by the photo editors I interviewed, . . . news coverage of poverty not only exaggerates the percentage of blacks among the poor but consistently portrays poor blacks more negatively than poor whites.”). Gilens specifically draws on the implicit bias literature to argue that “[s]ubconscious stereotypes represent the most plausible explanation for the discrepancy between the racial liberalism of news professionals and the negative images of blacks that are found in news coverage of poverty.” *Id.* at 153; see also Avery & Peffley, *supra* note 346, at 146–47 (arguing that news portrayal of welfare recipients as African American has decreased White support for welfare programs).

⁴⁰⁶ See Gilliam & Iyengar, *supra* note 6, at 562 (surveying crime statistics and reporting in Los Angeles County in 1996 and 1997).

⁴⁰⁷ See ENTMAN & ROJECKI, *supra* note 261, at 79 (reporting one study in which a plurality of White respondents guessed that Blacks constitute 60% of violent crime arrests, although the actual figure is closer to 40%).

⁴⁰⁸ See Diane M. Mackie et al., *Social Psychological Foundations of Stereotype Formation*, in *STEREOTYPES AND STEREOTYPING*, *supra* note 39, at 41, 50–51 (describing the phenomenon); see also Langevoort, *supra* note 21, at 1504 (noting the theme of “Illusory Correlations and Causation Biases” in behavioral law and economics).

⁴⁰⁹ See BROWN, *supra* note 50, at 86–88. The conventional explanation for this phenomenon is a general psychological bias in parsing rare events. See *id.* at 89–90.

⁴¹⁰ See Fiske, *supra* note 4, at 386 (explaining that negative information is informative, rare, and carries more weight).

⁴¹¹ See, e.g., FISKE & TAYLOR, *supra* note 26, at 373–74 (describing both “paired distinctiveness” and prior expectations as sources of illusory connections).

race, people nevertheless recall seeing schema-consistent minority perpetrators.⁴¹² We are not colorblind even when there is nothing to see.

Third, even if our recollections are accurate, our interpretations may be biased. Consider, for example, the “outgroup homogenization” effect. Outgroups are viewed as more homogeneous and monolithic than the ingroup,⁴¹³ and the ingroup is viewed as more heterogeneous than the average.⁴¹⁴ Therefore, even if the news conveys descriptively accurate information about the *mean* criminality of racial minority groups, the public still may seriously underestimate the *variance*. This would contribute to the fallacy of thinking that simply because 50% of crimes are committed by a group *X*, 50% of group *X* commit crimes. Consider how this tendency to view members of outgroups as monolithic could affect Arab Americans during our indefinite war on terror.

Another concern is the “fundamental attribution error” (FAE). The FAE is a general tendency to attribute the causes of behavior to dispositional, instead of situational, factors.⁴¹⁵ In other words, we tend to underweight contingent, environmental factors that cause a particular action and to highlight putatively stable factors such as personality traits instead.⁴¹⁶ This type of error is made more often about others and less often about ourselves. In other words, we each individually have the tendency to describe the causes of other people’s behavior as dispositional but to explain our own behavior as highly situational,

⁴¹² See Gilliam & Iyengar, *supra* note 6, at 564, 565 tbl.2 (reporting that in an experiment in which viewers saw no perpetrator, 63% recalled that they had *seen* one: 44% recalled a Black perpetrator, whereas only 19% recalled a White perpetrator).

⁴¹³ See BROWN, *supra* note 50, at 55 (noting studies); FISKE & TAYLOR, *supra* note 26, at 123 (describing the outgroup homogeneity effect). Not only are members of an outgroup seen as having less variability, but they are also conceptualized using fewer characteristics. *Id.*

⁴¹⁴ See, e.g., Fiske, *supra* note 4, at 367. The reasons for this bias are unclear. Some have suggested that greater familiarity with the ingroup leads to greater appreciation of individual differences and variety. However, experiments under minimal group paradigm conditions — conditions in which there was equal and near-zero information about both the ingroup and outgroup — suggest that familiarity is not the explanation. See BROWN, *supra* note 50, at 55–56. An alternative explanation suggests that the ingroup schema is more complex because it is more important (since it includes the perceiver), concrete (since it includes information about the perceiver), and provisional (since there is a greater motivation for accuracy). See *id.* at 56.

To be comprehensive, there is some evidence that ingroups sometimes view themselves as more homogeneous, such as minority groups that “close ranks” and find greater commonality. See *id.* at 57. In these circumstances, one might see a minority identity threatened by a majority identity, in which case the minority identity may perceive a homogeneity on identity-critical elements for the ingroup. See *id.*; *cf. id.* at 59 (suggesting that identity is more important for members of a minority and that ingroup members may strive to be more similar to the model ingroup member).

⁴¹⁵ See, e.g., FISKE & TAYLOR, *supra* note 26, at 67.

⁴¹⁶ See *id.* One source of this error may be that attribution to dispositional attributes is relatively automatic and easy to do, whereas identifying and evaluating the situational constraints require far more cognitive resources. See *id.* at 71–72.

since we naturally view ourselves as complicated, richly textured, and multidimensional.⁴¹⁷

The same disparity exists across social-category boundaries. Thomas Pettigrew calls this the “ultimate attribution error” (UAE) — “the tendency to accept the good for the ingroup and the bad for the outgroup as personal and dispositional, but more important, to explain away the bad for the ingroup and the good for the outgroup with situational attributions.”⁴¹⁸ Accordingly, when we see a Brown terrorist, we are inclined toward “outgroup essentialism” and interpret the violence as part of their way; by contrast, when we see John Walker Lindh or Timothy McVeigh, we see only wayward souls, saying nothing larger about our White selves.⁴¹⁹ In addition to being found in social cognition research, the FAE has been demonstrated in political science experiments employing the newscast paradigm.⁴²⁰ After a White mugshot, for example, participants emphasize societal variables in explaining the causes of crime; after a Black mugshot, participants emphasize individual nature.⁴²¹

Finally, our attitudinal interpretation of descriptively accurate information may vary tremendously and may not be especially amenable to “accurate” versus “inaccurate” classifications. Suppose the news reports that the average Jewish household savings rate is higher than the national average. Does this support the stereotype that Jews are stingy, or does it show that they are frugal? Could one successfully persuade the anti-Semite that the “stinginess” interpretation is somehow inaccurate, when he points to the accurate data as evidence? Or

⁴¹⁷ See *id.* at 72–74.

⁴¹⁸ Fiske, *supra* note 4, at 369 (citing Thomas F. Pettigrew, *The Ultimate Attribution Error: Extending Allport's Cognitive Analysis of Prejudice*, 5 PERSONALITY & SOC. PSYCHOL. BULL. 461 (1979)); see BROWN, *supra* note 50, at 101 (discussing evidentiary support comparing Hindus and Muslims in Donald M. Taylor & Vaishna Jaggi, *Ethnocentrism and Causal Attribution in a South Indian Context*, 5 J. CROSS-CULTURAL PSYCHOL. 162 (1974)); see also James D. Johnson et al., *Race, Media, and Violence: Differential Racial Effects of Exposure to Violent News Stories*, 19 BASIC & APPLIED SOC. PSYCHOL. 81, 86 (1997) (finding that attributions by a White audience of a hypothetical Black defendant's violent behavior were more dispositional than attributions of a hypothetical White defendant's violent behavior). This sort of ingroup advantage in interpretation is reflected nicely in LIB measurements: A narrative describes some action made by an individual. Then, participants are told to describe what the individual did. When the actor is in the ingroup, the action is described in concrete terms, for example, “punching someone.” By contrast, when the actor is in the outgroup, the action is described as “aggressive.” See Fiske, *supra* note 4, at 370. Note the difference in the level of abstraction in description.

⁴¹⁹ In the terrorism context, this disparate treatment has been noted by others. See, e.g., Thomas W. Joo, *Presumed Disloyal: Executive Power, Judicial Deference, and the Construction of Race Before and After September 11*, 34 COLUM. HUM. RTS. L. REV. 1, 42 (2002); Leti Volpp, *The Citizen and the Terrorist*, 49 UCLA L. REV. 1575, 1584–85 (2002).

⁴²⁰ See SHANTO IYENGAR, IS ANYONE RESPONSIBLE?: HOW TELEVISION FRAMES POLITICAL ISSUES 43–45 (1991) (finding that the race of the suspect in a crime story had more of an effect on viewers in assigning blame than the way the crime was framed).

⁴²¹ See *id.*

suppose the local news shows Korean store owners defending their shops with guns during looting and rioting. Does this support the stereotype that Koreans are self-reliant, or does it show that they are vigilantes?⁴²²

The accuracy objection has lost much of its force. The data are likely unrepresentative. Even when they are not, the biases of illusory correlation, outgroup essentialism, and outgroup homogenization — each of which maps nicely to themes in critical race theory — conspire to produce inaccurate assessments. They incline us to see more minority-committed crimes than actually exist, to attribute those bad acts to stable, fixed dispositions (either bad genes or stable pathological culture),⁴²³ and to think that racial minorities are monolithically “this” way. Of course, our cognitive biases cannot be attributed to the local news itself. However, only formalism would allow us to bank on the “accuracy” of the local news and to refuse to take into account the empirical findings of our cognitive limitations.

D. Virus Protection

The social cognition studies that I have presented are not without their ambiguity, confusion, and contradiction. They often raise as many questions as they answer. That said, a prima facie case has been made about the existence of implicit bias, its dissociation from explicit self-reports of bias, its measurability through reaction time designs, and its impact on behavior. Although weaker, a prima facie case has also been made that a nontrivial stream of negative meanings is provided through the local news. These images not only strengthen long-term, well-learned associations between certain racial categories and certain racial meanings, but also activate specific responses or states. Social scientists will, I believe, further confirm these claims over the next decade. What then? How might we address this injustice or inefficiency?⁴²⁴

⁴²² For a sharp critique by a veteran Korean American journalist and activist of how the media catalyzed the 1992 riots in Los Angeles and the targeting of Koreans for violence, see K.W. Lee, *Sa-I-Gu: A Haunting Prelude to the Fire Next Time*, at <http://www.angelfire.com/rant/tkd/> (last visited Feb. 13, 2005).

⁴²³ This essentialist attribution may be encouraged by another bias sometimes called “hedonic relevance.” When we are impacted negatively by another person’s actions, we tend to favor dispositional attributions over situational ones. See FISKE & TAYLOR, *supra* note 26, at 32.

⁴²⁴ It is standard operating procedure for economists to characterize the bad behavioral consequences of consuming information products as a negative externality that prompts overproduction. See, e.g., James T. Hamilton, *Private Interests in “Public Interest” Programming: An Economic Assessment of Broadcaster Incentives*, 45 DUKE L.J. 1177, 1178 (1996) (justifying governmental intervention as a remedy for the underproduction of programming that produces positive externalities and for the overproduction of programming that produces negative externalities).

1. *Recoding the Public Interest.* — Recall how we got here. Given our First Amendment, broadcast was the appropriate communications medium on which to focus our inquiry into implicit bias. The touchstone of regulating broadcast is the “public interest.”⁴²⁵ That standard has been recently rearticulated in a manner that perversely encourages production of Trojan Horse viruses.⁴²⁶ At the least, we should contest this definition of the public interest.

First, we should reject the strong linkage the FCC made between the public interest and the number of hours of local news aired. In its Media Ownership Order, the FCC lauded more local news as unequivocally furthering the public interest.⁴²⁷ Nowhere did the FCC consider the costs of local news, as measured by political scientists and social cognitionists. That was a mistake. In any future Notice of Proposed Rulemaking, a public comment could flag this omission simply by attaching a copy of this Article as support. If the FCC then completely ignores the local news-implicit bias link in its analysis, a court could find its decisionmaking arbitrary and capricious, in violation of the Administrative Procedure Act.⁴²⁸ At the least, the Commission would have to give reasons for dismissing these arguments.

Second, the FCC should reconsider its decision to limit viewpoint diversity analysis to news and public affairs programming. Recall that various stakeholders, such as the Fox Network, wanted counterstereotypic entertainment programming to count in the viewpoint diversity calculus.⁴²⁹ The FCC declined. But if we care about implicit bias, counting only local news in the public interest analysis is perverse. In the malleability studies, for instance, many of the positive minority images that decreased implicit bias were entertainment celebrities.⁴³⁰ In other words, the best scientific evidence is that repeated exposure to Bill Cosby, no doubt in part because he is also “Dr. Huxtable,” decreases our implicit bias. Of course, there may be substantial costs to opening this diversity can of worms. But at the very least, the FCC should be forced to make a public accounting.

This would be no easy task for the Commission. Consider again the popular sitcom *The Cosby Show*. On the one hand, some have criticized the show for refusing to depict the reality of racial discrimination, for displaying an upper-middle-class family that just happens

⁴²⁵ See *supra* p. 1545.

⁴²⁶ See *supra* pp. 1549–50.

⁴²⁷ See *supra* p. 1495.

⁴²⁸ See 5 U.S.C. § 706 (2000).

⁴²⁹ See *supra* p. 1547. To be sure, the shows identified by the Fox Network may have been entirely self-serving. The networks probably also broadcast entertainment shows that would clearly increase implicit bias: think reality crime shows.

⁴³⁰ See *supra* pp. 1557–58 & note 366.

to be Black.⁴³¹ Critics claim that this portrayal sugarcoats African American reality, even at that tax bracket. The lesson taught and learned might be that the Huxtables can succeed in America; why can't the rest of "them"?⁴³² Arguably, this could exacerbate the type of *explicit* bias measured by, say, the ambivalent racism scale, which gauges resentment of Blacks on the belief that they have not taken full advantage of the opportunities afforded them.⁴³³ On the other hand, many applauded the show for providing positive images of the Black family.⁴³⁴ And here is how social cognition radically disrupts left-of-center conventional wisdom: watching *The Cosby Show* probably decreases *implicit* bias. How these two dissociated processes interrelate and superimpose we do not yet understand.

If *The Cosby Show* reveals the tension between explicit and implicit bias, then the wildly popular show *Queer Eye for the Straight Guy* reveals the tension between (cognitive) stereotypes and (affective) prejudice. Consider what a study might soon reveal.⁴³⁵ On the one hand, viewers' stereotypes about gay men may strengthen after watching this show, as the five characters "gay it up" for entertainment value. But, and here is the fascinating complication, viewers' affect toward gay men could become more positive; in other words, their negative evaluations could decrease. Whether *Queer Eye* is good or bad thus depends partly on the context. For example, if the goal is stopping hate crimes, then the show seems a net positive since it makes us feel more positive toward gay men. By contrast, if the goal is to break through glass ceilings and walls for gay men pursuing professions seen as aggressively masculine, then the show may be a net negative.

All this remains speculation. Yet lack of perfect understanding does not mean we should ignore the social science, but rather that we

⁴³¹ See, e.g., ENTMAN & ROJECKI, *supra* note 261, at 71 (arguing that *The Cosby Show* suggests that racial discrimination has disappeared); RICHARD JACKSON HARRIS, A COGNITIVE PSYCHOLOGY OF MASS COMMUNICATION 70 (4th ed. 2004) (noting the danger that *The Cosby Show* will be seen as evidence that affirmative action is no longer necessary).

⁴³² Of course, the other complaint is that Blacks appearing in entertainment programming are stereotypical. See, e.g., Thomas E. Ford, *Effects of Stereotypical Television Portrayals of African-Americans on Person Perception*, 60 SOC. PSYCHOL. Q. 266, 267 (1997) (suggesting that African-American characters are "frequently based on disparaging stereotypes" and that such portrayals have "traditionally been more pronounced in programs with predominantly African-American casts").

⁴³³ See *supra* note 78.

⁴³⁴ See, e.g., Leonard M. Baynes, *White Out: The Absence and Stereotyping of People of Color by the Broadcast Networks in Prime Time Entertainment Programming*, 45 ARIZ. L. REV. 293, 314 (2003) (characterizing the show positively and pointing out that NBC consulted with Harvard psychiatrist Dr. Alvin Poussaint on scripts and characters).

⁴³⁵ Undergraduates in a psychology class of Professor Denise Sekaquaptewa at the University of Michigan have done some preliminary work on this topic. No publishable results exist yet. E-mail from Denise Sekaquaptewa to Jerry Kang (Mar. 22, 2004, 10:39 PM) (on file with the Harvard Law School Library).

should investigate it further. In social psychological terms, the FCC's "attitude" toward local news must grow more complicated. Instead of blind enthusiasm for local news, there should be ambivalence.

Third, further study through a Notice of Inquiry is warranted. Relevant lines of inquiry include:

- How should the "public interest" be defined?
- What are the costs and benefits of using "local news" to define the "public interest"?
- How might the quality of "local news" be measured?
- Are there broadcast practices or guidelines that might mitigate the implicit bias increased by viewing violent crime on local news?
- How might the quantity and quality of public affairs programming be measured?
- How might viewpoint diversity in entertainment programming be measured?

Such an inquiry could catalyze a more thorough conversation about what the "public interest" ought to mean and how it should be operationalized.⁴³⁶ In this discussion, we must face the difficult question of what sort of content might replace local news if news is less encouraged by regulators. There is reason to think that advertisements,⁴³⁷ dramas, reality shows, and sitcoms⁴³⁸ also function as Trojan Horses.

⁴³⁶ I concede that this may open a can of regulatory worms, with the end result being a louder, more insistent call against broadcast exceptionalism — to deny the FCC any content control over broadcast. I am willing to take this risk in order to push the Commission to reconsider its current conceptualization of the public interest.

⁴³⁷ See generally ENTMAN & ROJECKI, *supra* note 261, at 162–81; Deseriee A. Kennedy, *Marketing Goods, Marketing Images: The Impact of Advertising on Race*, 32 ARIZ. ST. L.J. 615, 638–60 (2000) (providing a history of race and advertising, with a focus on African Americans). Kennedy suggests as a response to her findings a ban on discriminatory commercial content. *Id.* at 672–73. On a related point, some commentators have alleged that advertisers discriminate against minority-owned/minority-audience radio stations. See, e.g., Robert Millar, *Racism is in the Air: The FCC's Mandate to Protect Minorities from Getting Shortchanged by Advertisers*, 8 COMMLAW CONCEPTUS 311, 311–13 (2000).

⁴³⁸ See generally ENTMAN & ROJECKI, *supra* note 261, at 144–61 (discussing race relations on primetime television). In one study, Thomas Ford exposed forty participants to either neutral or stereotypical comedy skits about African Americans, with the latter skits depicting them as "poor, uneducated, and prone to acts of crime and violence." Ford, *supra* note 432, at 269. Next, participants, who were led to believe that they were participating in an entirely different study, read an ambiguous narrative about a target, named either Tyrone or Todd, accused of assaulting a fellow student. *Id.* To verify that the names "Tyrone" and "Todd" mapped to Black and White respectively, Ford sampled 45 students from the same population and asked them to map 25 names, including Tyrone and Todd. Forty-five out of 45 mapped Tyrone to African American. *Id.* Forty-four out of 45 mapped Todd to White (non-Jewish); 1 mapped Todd to Jewish. *Id.* The stereotypical comedy skit had no impact on the guilt ratings of Todd, that is, the participant's impression of Todd's likely guilt and the strength of the case against him. *Id.* at 270–71. By contrast, the stereotypical skit produced higher guilt ratings for Tyrone. *Id.* at 271 ($p = 0.024$). Perhaps the most critical comparison is whether participants judged Todd or Tyrone differently in

If local news is replaced, for example, by reality cop shows, the impact on implicit bias against racial minorities might be worse. The same may be true if the replacement is a grotesquely stereotypical sitcom. That said, one cannot imagine a station pointing to such content as uncontroversial evidence of its serving the “public interest,” at least not with a straight face.

Fourth, the FCC in conjunction with media elites should publicly explore how the news exacerbates implicit bias, with an eye toward voluntary development and adoption of “best journalistic practices.” Examples include scrupulously checking against disparate treatment of minority suspects in crime stories, minimizing unnecessary racial mapping, and avoiding the worst inflammatory images.⁴³⁹ These best practices could extend beyond crime stories, to seek more diverse representation of “experts” and to emphasize the value of positive stories of racial minorities promoting safety and harmony within the local community.⁴⁴⁰ The FCC could catalyze this conversation through various informal bully pulpit and jawboning techniques. In addition, the FCC could institute greater self-monitoring and self-reporting requirements about the percentage of news minutes focusing on violent crime during some randomly sampled time periods.⁴⁴¹ Such data could bring social and market pressure to bear on how stations discharge their public interest responsibilities.

In the vast electronic ocean of vicarious experiences swirling around us, who knows what total impact crime stories in local news have in comparison to representations of minorities in music videos,⁴⁴²

terms of guilt after being exposed to the stereotypical skit. Although the mean guilt ratings were higher for Tyrone than for Todd after exposure, that difference did not reach statistical significance. *See id.* ($p = 0.07$).

⁴³⁹ There are reasons to think that such efforts may have only limited effect. For example, in one experiment, participants read violent news stories without mention of race. This colorblind prime nevertheless prompted Black and White participants alike to alter their interpretations of a Black target’s violent behavior to be more dispositional than situational. No such impact was seen for the White target. Johnson et al., *supra* note 418, at 86.

⁴⁴⁰ *See generally* AV WESTIN, BEST PRACTICES FOR TELEVISION JOURNALISTS 21–26 (discussing bias in newsrooms). Westin provides an anonymous quotation from a newsworker that makes the point well:

We were working on a story about whether parents or peers have more influence on kids. We were showing kids playing games in a video arcade, and the cameraman shot some general cover footage including an image of a little black boy with a gun. In the edit room, the editor picked it. I flinched when I saw it. My producer immediately realized why when she saw the look on my face. She said, “Black kid holding a gun, it’s a stereotype.” She changed the shot. We see so many African-American boys with guns in the newspaper and magazines. Every black boy does not own a gun.

Id. at 25.

⁴⁴¹ *See, e.g.*, Hamilton, *supra* note 424, at 1187 (suggesting that electronic access to children’s television compliance records would make public monitoring easier).

⁴⁴² Rudman and Lee found that a thirteen-minute rap music prime can activate “Black = hostile” schemas, which influence the interpretation of hostility in ambiguous actions by Black tar-

video games,⁴⁴³ entertainment programming,⁴⁴⁴ and motion pictures? However, regardless of the relative significance of news, we should not allow a poor articulation of the “public interest” standard to go unchallenged simply because the problems of negative stereotypes and prejudice against racial minorities are so enormous. Finally, although the focus has been on the local news, this public discussion would shine a new light on racial meanings generated and delivered throughout all media. Maybe nothing will be done about it, in the name of profit and freedom of expression. But at the least, we as a society will better understand what we have chosen to do, through act and omission.

2. *Thought Experiments.* — I now take a more radical turn, by engaging in two thought experiments. At the outset, I concede that the scientific case for the efficacy of these proposals may not be strong, depending on where the burden of proof is set. Nonetheless, considering more provocative measures may be illuminating. In pursuing these thought experiments, I am self-consciously taking advantage of the different craft norms within disciplines, which allow legal analysts to ask questions that scientists would typically not.

So I return to the metaphor of local news as Trojan Horse viruses. In these terms, the prior recommendation to recode the “public interest” standard was a call to stop encouraging the production of programs that turned out to be Trojan Horses. But in the realm of computer security, more aggressive antivirus strategies are available: build a firewall to decrease the exposure, and push out disinfectants to treat the infection.

(a) *Firewall: Capping Crime Stories.* — Questioning governmental encouragement of the Trojan Horses of race is one thing. But might we go a step further and affirmatively build a firewall against them? We ban obscenity outright. It would, however, be inconsistent with any reasonable interpretation of the First Amendment to try to ban local news, crime stories, or even particular ways in which stories are conveyed.⁴⁴⁵ Also, the technique of late-night “time channeling” em-

gets. Rudman & Lee, *supra* note 361, at 136, 138–39, 141–42. Music videos, with their often spectacular visuals, may well have even greater effect.

⁴⁴³ See, e.g., Michel Marriott, *The Color of Mayhem*, N.Y. TIMES, Aug. 12, 2004, at G1 (relaying criticism that games such as *Grand Theft Auto: San Andreas* provide stereotypical “images of black youths committing and reveling in violent street crime”).

⁴⁴⁴ For an interesting analysis of racial stereotypes in entertainment programming, see Baynes, *supra* note 434, at 313–30. Baynes argues that current media practices of omitting minorities or presenting them stereotypically can be legally actionable under 47 U.S.C. § 151, which mandates the making of communication available “without discrimination on the basis of race, color, religion, national origin, or sex,” and § 309(k), which addresses renewal of license. *Id.* at 332–36. He further argues that such action would be analogous to enforcement of the ban on discriminatory advertisements under § 804(c) of the federal Fair Housing Act. *Id.* at 342–49.

⁴⁴⁵ Voluntary decisions not to mention race unless relevant to the story are another matter, and may reflect good journalistic judgment. Although one may be able to do this in text, since names

ployed against indecency seems terribly underinclusive because the problem of implicit bias does not affect kids alone.⁴⁴⁶ Finally, the V-chip in its current incarnation is irrelevant. The V-chip helps parents protect children (and themselves) from exposure to coarse language, sex and nudity, and violence — content that they do not want. We do, however, want news — which, by the way, is not rated within the V-chip system.

What about simply throttling the supply (regardless of demand)? For example, the FCC has historically limited the number of minutes of commercials aired per hour.⁴⁴⁷ Currently, about 25% of news reporting is about crime;⁴⁴⁸ so, suppose that the FCC suggests a 15% cap.⁴⁴⁹ It could justify the cap by invoking the government interest in decreasing implicit bias against racial minorities. The 15% cap would be a guideline, not a quota, the meeting of which would become part of the record relevant to license renewal, a process that is largely perfunctory. Would such a guideline violate the First Amendment?⁴⁵⁰

It would be content-based regulation, thus warranting the formal application of strict scrutiny.⁴⁵¹ Therefore, the government would have to identify a “compelling interest” and also demonstrate that the techniques it has adopted to further this interest are “narrowly tailored.” But this standard of review is warped by the communications

may not guarantee racial mapping, one cannot so easily do this with images. *See* Kang, *supra* note 19, at 1156 (discussing different techniques of racial mapping as a function of text, voice, and video).

⁴⁴⁶ If children were found to be particularly amenable to certain forms of racial programming, time channeling could be more easily justified. For a discussion of how children develop a theory of race, see LAWRENCE A. HIRSCHFELD, *RACE IN THE MAKING: COGNITION, CULTURE, AND THE CHILD’S CONSTRUCTION OF HUMAN KINDS* 83–119 (1996).

⁴⁴⁷ *See, e.g.*, TV Deregulation Order, *supra* note 285, at 1101–05 (describing commercialization requirements, then repealing them). In 1973, the FCC adopted a sixteen-minute-per-hour commercialization guideline. *See id.* at 1101.

⁴⁴⁸ *See supra* p. 1549.

⁴⁴⁹ Obviously, one could intervene with greater precision, for example by targeting only those crime stories featuring racial minority suspects. But such a scalpel-like approach seems much more like viewpoint discrimination, which raises far greater constitutional concerns.

⁴⁵⁰ A prior threshold question is whether the FCC has power to issue such a regulation without specific statutory authorization. *Cf.* Motion Picture Ass’n of Am. v. FCC, 309 F.3d 796, 807 (D.C. Cir. 2002) (reversing and vacating the Commission’s order requiring video description for the blind for lack of statutory authorization). To keep focused on the thought experiment, please suppose that Congress provides whatever specific authorization deemed necessary by the judiciary.

⁴⁵¹ There is no reasonable argument to suggest that this regulation is content-neutral since it is designed specifically to go after a particular type of communicative impact. Although one could imagine a First Amendment that distinguishes between communicative impact mediated through conscious processes as compared to unconscious ones, our First Amendment does not do so. *See* Frederick Schauer, *Speech and “Speech” — Obscenity and “Obscenity”: An Exercise in the Interpretation of Constitutional Language*, 67 GEO. L.J. 899, 921 (1979).

medium.⁴⁵² As the Supreme Court taught in *FCC v. Pacifica Foundation*,⁴⁵³ “the broadcast media have established a uniquely pervasive presence in the lives of all Americans.”⁴⁵⁴ Moreover, broadcast is “uniquely accessible to children, even those too young to read.”⁴⁵⁵ These distinguishing features, as well as the extensive history of broadcast regulation, have justified greater content regulation of broadcast. In addition, the scarcity/interference rationale justifies licensing and other “public interest” requirements.⁴⁵⁶ In sum, although “strict scrutiny” will be applied, in the broadcast context the law is de facto “strict scrutiny light.”

Whether the state’s interest in decreasing implicit bias can be considered “compelling” depends on both empirical and normative judgments. The empirical question asks whether the problem identified (implicit bias exacerbated by local news crime reports) actually exists or is mere conjecture.⁴⁵⁷ The subsequent normative question is whether the behavioral consequences of implicit bias, once empirically verified, should count as “compelling.” In answering these questions, we can be guided by way of analogy — the most illuminating one being broadcast indecency.

In 1992, Congress instructed the FCC to adopt time-channeling regulations on indecent programming.⁴⁵⁸ The D.C. Circuit reviewed these regulations under a strict-scrutiny-light algorithm appropriate for broadcast.⁴⁵⁹ In analyzing the compelling interest prong of the means-ends analysis, the court relied on language from *New York v. Ferber*:⁴⁶⁰

⁴⁵² Consider, for instance, the longstanding legality of the fairness doctrine, which was voluntarily interred by the FCC. That decision was affirmed by the courts on “public interest” grounds, not on some notion of First Amendment necessity. See *Syracuse Peace Council v. FCC*, 867 F.2d 654, 669 (D.C. Cir. 1989); see also 47 U.S.C. § 312(a)(7) (2000) (requiring minimum access to advertising time of federal candidates); 47 U.S.C. § 315(a) (requiring equal access for political candidates).

⁴⁵³ 438 U.S. 726 (1978).

⁴⁵⁴ *Id.* at 748.

⁴⁵⁵ *Id.* at 749.

⁴⁵⁶ See, e.g., *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 390 (1969) (upholding fairness doctrine, which created right-of-reply under certain circumstances).

⁴⁵⁷ See, e.g., *Turner Broad. Sys., Inc. v. FCC*, 512 U.S. 622, 664 (1994) (“When the Government defends a regulation on speech as a means to . . . prevent anticipated harms, it . . . must demonstrate that the recited harms are real, not merely conjectural, and that the regulation will in fact alleviate these harms in a direct and material way.”).

⁴⁵⁸ See Public Telecommunications Act of 1992, Pub. L. No. 102-356, § 16(a), 106 Stat. 949, 954 (codified at 47 U.S.C. § 303 note (2000)).

⁴⁵⁹ See *Action for Children’s Television v. FCC*, 58 F.3d 654, 660 (D.C. Cir. 1995) (“While we apply strict scrutiny to regulations of this kind regardless of the medium affected by them, our assessment of whether section 16(a) survives that scrutiny must necessarily take into account the unique context of the broadcast medium.”).

⁴⁶⁰ 458 U.S. 747 (1982).

It is evident beyond the need for elaboration that a State's interest in safeguarding the physical and psychological well-being of a minor is compelling. A democratic society rests, for its continuance, upon the healthy, well-rounded growth of young people into full maturity as citizens. Accordingly, we have sustained legislation aimed at protecting the physical and emotional well-being of youth even when the laws have operated in the sensitive area of constitutionally protected rights.⁴⁶¹

When petitioners pointed out that "no causal nexus ha[d] been established between broadcast indecency and any physical or psychological harm to minors,"⁴⁶² the court rejected the idea that there had to be any "clinically measurable injury."⁴⁶³ The full explanation is worth quoting:

[T]he Supreme Court has never suggested that a scientific demonstration of psychological harm is required in order to establish the constitutionality of measures protecting minors from exposure to indecent speech. In *Ginsberg*, the Court considered a New York State statute forbidding the sale to minors under the age of 17 of literature displaying nudity even where such literature was "not obscene for adults" The Court observed that while it was "very doubtful" that the legislative finding that such literature impaired "the ethical and moral development of our youth" was based on "accepted scientific fact," a causal link between them "had not been disproved either." The Court then stated that it "d[id] not demand of legislatures scientifically certain criteria of legislation. . . ."

. . . .

[In addition,] the Supreme Court has recognized that the Government's interest in protecting children extends beyond shielding them from physical and psychological harm. The statute that the Court found constitutional in *Ginsberg* sought to protect children from exposure to materials that would "impair[] [their] *ethical and moral* development."

. . . .

Congress does not need the testimony of psychiatrists and social scientists in order to take note of the coarsening of impressionable minds that can result from a persistent exposure to sexually explicit material just this side of legal obscenity. The Supreme Court has reminded us that society has an interest not only in the health of its youth, but also in its quality. As Irving Kristol has observed, it follows "from the proposition that democracy is a form of self-government, . . . that if you want it to be a meritorious polity, you have to care about what kind of people govern it."⁴⁶⁴

⁴⁶¹ *Action for Children's Television*, 58 F.3d at 661 (quoting *Ferber*, 458 U.S. at 756–57).

⁴⁶² *Id.*

⁴⁶³ *Id.* at 663.

⁴⁶⁴ *Id.* (second, fourth, and fifth alterations in original) (emphasis added) (citations omitted) (quoting *Ginsberg v. New York*, 390 U.S. 629, 634, 641–43 (1968); and IRVING KRISTOL, ON THE DEMOCRATIC IDEA IN AMERICA 41–42 (1972)).

On the empirical question whether sexually explicit television harmed minors, the court dismissed any need for social science evidence. Irving Kristol's musings sufficed. If we do not need social psychological evidence about the harm of indecent television, why should we demand social psychological evidence about the harm of implicit bias-inducing programming? There is the obvious distinction based on age. Broadcast indecency policy is targeted at raising healthy and virtuous youth; implicit bias policy is targeted at everyone, including adults whom we are chary of making virtuous through content regulation. But this distinction is not as sharp as it may first seem. For instance, is the above statement really true? Is our current indecency policy targeted solely at protecting children, or is it part of a broader cultural stance of what is appropriate for adults as well? And if the focus on children is crucial for legal palatability, we could design a time-channeled firewall strategy tailored to protect kids the most. This may sound crassly instrumental, but as developmental psychologists cooperate with social cognitionists, we may eventually learn that implicit biases accrete most powerfully during childhood, before critical cognitive faculties have fully matured. In other words, while our explicit racial meanings may reflect self-critically chosen beliefs as adults, implicit racial meanings have simply attached like barnacles since our infancy.⁴⁶⁵

Maybe there is simply greater consensus on the immorality of "indecency" as compared to the immorality of "implicit bias." But I do not believe that indecency raises uncontroversial questions about basic manners and civility (therefore needing no scientific evidence), whereas implicit bias is a patently obvious ideological weapon of political correctness (therefore needing scientific evidence). Recall Justice Brennan's fierce dissent in *Pacifica*, which addressed George Carlin's "Filthy Words" monologue. There he pointed out that "some parents may actually find Mr. Carlin's unabashed attitude towards the seven 'dirty words' healthy."⁴⁶⁶ Brennan showed a deep concern about legally codifying a tyranny of majoritarian prudishness:

[I]n our land of cultural pluralism, there are many who think, act, and talk differently from the Members of this Court, and who do not share their fragile sensibilities. It is only an acute ethnocentric myopia that enables the Court to approve the censorship of communications solely because of the words they contain.⁴⁶⁷

⁴⁶⁵ See *supra* note 81 (describing Devine's model). Dovidio explains how implicit and explicit attitudes reflect "dual attitudes," which arise developmentally. Original attitudes from our youth may not be replaced as much as they are submerged by newer, conscious, explicitly chosen attitudes. See Dovidio, *supra* note 110, at 839.

⁴⁶⁶ *FCC v. Pacifica Found.*, 438 U.S. 726, 770 (1978) (Brennan, J., dissenting).

⁴⁶⁷ *Id.* at 775.

Finally, one cannot suggest that there is greater scientific consensus on harms caused by indecency as compared to the harms caused by implicitly biasing news reports. After all, if such a consensus existed, the Court would have simply relied on it to find a compelling interest. In sum, given that the Court has demanded essentially no empirical foundation in the comparable case of indecency, one can reasonably argue that no case need be made as to implicit bias. Even if such a case must be made, I think that the social science can provide one. Finally, if child welfare must lie at the core of any such firewall attempts, the strategy could be tailored to protect kids first, with incidental impact on adults.

After answering the empirical question, we must also make a final normative judgment. In other words, even if implicit bias does what I claim it does, is that important enough to count as “compelling”? First Amendment doctrine does not provide a clear test to answer such a question. So we reason through close analogy to other government interests that have been deemed “compelling.” We know, for example, that “safeguarding the physical and psychological well-being” of children is a compelling interest.⁴⁶⁸ Drawing on equal protection case law, we know that remedying racial discrimination as well as pursuing educational diversity count as compelling interests. In the end, the government interest in decreasing implicit bias should be deemed “compelling” as well.

Few speech regulations are struck down on the first prong of compelling interest. Rather, what does them in is their inadequate tailoring. After all, no law is ever perfectly matched to a particular goal. Overinclusiveness and underinclusiveness are inevitable, and a skeptical court can easily raise the possibility of some other solution that could have done a better job by tweaking different technologies, social norms, and legal regulations.

Would a regulation suggesting an upper guideline of 15% crime stories in local news be “narrowly tailored” to the goal of decreasing implicit bias caused by viewing crime stories? Phrased this way, the answer is “yes” — but tautologically so. Framed another way, however, the strategy would clearly *not* be narrowly tailored. If we ask specifically whether limiting crime stories in local news is narrowly tailored to the goal of ending *all racial bias*, the answer is “no”: we know that this intervention will have no such grandiose impact. But this latter formulation of the inquiry cannot be correct, for such an inquiry

⁴⁶⁸ *New York v. Ferber*, 458 U.S. 747, 756–57 (1982) (quoting *Globe Newspaper Co. v. Superior Court*, 457 U.S. 596, 607 (1982)).

would doom all attempts at content control. After all, no solution is ever that total in effect.⁴⁶⁹

A more meaningful way to ask the tailoring question is to recognize the tradeoff between the good impact of crime news (as information about newsworthy events) and the bad impact of crime news (as Trojan Horse). In this tradeoff, what is the evidence that 15% is the right line to draw? Framed this way, the indecency cases again provide the benchmark for comparison.

In the indecency context, the tradeoff was between the good impact of adult access to potentially indecent content (basic First Amendment access) and the bad impact of minors' access to the same (corruption of minors). In *Action for Children's Television v. FCC*, the D.C. Circuit had to decide whether a midnight-to-6 a.m. safe harbor was the right line to draw. In its analysis, the court was remarkably deferential to the hours chosen by the FCC and Congress to time-channel indecent content. After looking at the numbers of adults versus children watching television in the late hours, the court concluded that a midnight-to-6 a.m. safe harbor was justified.⁴⁷⁰ It made clear that extraordinary "fine-tuning" was not necessary:

It is, of course, the ultimate prerogative of the judiciary to determine whether an act of Congress is consistent with the Constitution. Nevertheless, we believe that deciding where along the bell curves of declining adult and child audiences it is most reasonable to permit indecent broadcasts is the kind of judgment that is better left to Congress, so long as there is evidence to support the legislative judgment.⁴⁷¹

This kind of reasoning, if consistently applied to the crime news cap — admittedly, a big "if" — bodes well for its constitutionality.

In comparing indecency regulations to the hypothetical firewall of limiting crime stories, I am not arguing that precedent approving the former requires approving the latter. Any competent legal analyst can distinguish the fact patterns and case law. Rather, my modest point is that a consistent application of broadcast indecency doctrine makes constitutionally *plausible* a soft cap on crime stories in local news, especially when kids are in the audience in substantial numbers. A well-designed, administrable regulation, applied sensibly to include, for instance, generous exceptions for especially newsworthy events such as celebrity trials, could be upheld. If the social science in the next decade suggests that such a cap would substantially decrease implicit bias

⁴⁶⁹ Cf. *United States v. Am. Library Ass'n*, 539 U.S. 194, 199 (2003) (upholding the Children's Internet Protection Act (CIPA)). If CIPA had to protect children from all pornography, it would have been deemed underinclusive and therefore not narrowly tailored — but it was upheld.

⁴⁷⁰ *Action for Children's Television v. FCC*, 58 F.3d 654, 656 (D.C. Cir. 1995).

⁴⁷¹ *Id.* at 667.

— another big “if” — then we should be open to having this debate without constitutional concerns’ short-circuiting the conversation.

(b) *Disinfection: Public Service Announcements.* — The other antivirus strategy is disinfection, to push out antidotes to the Trojan Horses that we admit. For a timely example of such a strategy, consider the Broadcast Decency Enforcement Act of 2004,⁴⁷² recently passed in the House, which increases penalties for broadcast indecency in the wake of the Howard Stern and Janet Jackson incidents. The Act would authorize the FCC to

impos[e] a penalty under this section, [and] require the licensee or permittee to broadcast public service announcements that serve the educational and informational needs of children. Such announcements may be required to reach an audience that is up to 5 times the size of the audience that is estimated to have been reached by the obscene, indecent, or profane material, as determined in accordance with regulations prescribed by the Commission.⁴⁷³

What is interesting is the use of educational programming to counter the exposure to indecency, as if somehow ads against smoking counter Janet Jackson’s wardrobe malfunction.

In more familiar doctrinal terms, disinfection is counterspeech. And if the firewall approach felt uncomfortably like censorship, disinfection avoids such associations. To be clear, disinfection does not necessarily take the form of ponderous documentaries about race. Although such shows may decrease explicit bias, they may not be best suited to tweak implicit bias. As John Bargh said, we must “fight[] automatic fire with automatic fire.”⁴⁷⁴ A food analogy may be helpful. People will eat food (by comparison, watch local news) because they need it to survive. But eating food causes tooth decay (just as watching news increases implicit bias). Instead of only cajoling people to brush their teeth (that is, show PBS documentaries during Black History Month), why not also fluoridate the water (in other words, add positive counterspeech)?⁴⁷⁵

⁴⁷² H.R. 3717, 108th Cong. (2004).

⁴⁷³ *Id.* at § 6(c).

⁴⁷⁴ See Blasi, *supra* note 74, at 1254 (quoting John A. Bargh, Presentation at the UCLA Department of Psychology (Jan. 25, 2002)) (internal quotation marks omitted).

⁴⁷⁵ I am aware of the controversies, both historical and present, of fluoridating the public water supplies. See, e.g., John Remington Graham & Pierre-Jean Morin, *Highlights in North American Litigation During the Twentieth Century on Artificial Fluoridation of Public Water Supplies*, 14 J. LAND USE & ENVTL. L. 195, 198–200 (1999) (suggesting that fluoride is a carcinogen and mutagen and challenging whether states have constitutional authority to fluoridate public water supplies); Harold P. Green, *The Law-Science Interface in Public Policy Decisionmaking*, 51 OHIO ST. L.J. 375, 403 (1990) (pointing out that “when fluoridation of public water supplies was under consideration as a fix for dental caries in children, there is no record that any serious thought was given to legislation mandating periodic visits to the dentist, or restricting the manufacture and sale of candy and chewing gum”); see also *Quiles v. City of Boynton Beach*, 802 So. 2d 397, 399

The malleability studies suggest that positive images of racial minorities alter the cache of racial meanings as well as make positive exemplars more accessible. So, consider numerous variations on a strategy of debiasing public service announcements (d-PSAs).⁴⁷⁶ For purposes of argument, suppose that social cognitionists confirm that d-PSAs decrease implicit bias in substantial amounts.⁴⁷⁷ Even if the effect is temporary, viewers would be debiased daily, given the amount of television that Americans watch.⁴⁷⁸ How might we utilize d-PSAs? The strategy could differ along the following variables: state action (mandatory/voluntary); notice (subliminal/supraliminal); and consent (opt-out/opt-in).

First, consider state action. On the one hand, the FCC could require broadcast licensees to show some quota of d-PSAs. This would be state action that burdened licensee speech. That is not to say that stations would challenge such a regulation or that they would succeed

(Fla. Dist. Ct. App. 2001) (holding that fluoridation does not violate a citizen's right to refuse medical treatment); *Safe Water Ass'n v. City of Fond du Lac*, 516 N.W.2d 13, 15 (Wis. Ct. App. 1994) (upholding city council's fluoridation program as a valid exercise of state police power); David L. Shapiro, *Courts, Legislatures, and Paternalism*, 74 VA. L. REV. 519, 526, 547-48 (1988) (raising paternalism critique). Nonetheless, the point is that we do take action notwithstanding such controversy. Of course, atoms in our bodies (fluoride) are viewed as qualitatively different from bits in our brains (PSAs).

⁴⁷⁶ A PSA is

any announcement (including network) for which no charge is made and which promotes programs, activities, or services of federal, state, or local governments (e.g., recruiting, sale of bonds, etc.) or the programs, activities or services of nonprofit organizations (e.g., United Way, Red Cross blood donations, etc.) and other announcements regarded as serving community interests, excluding time signals, routine weather announcements, and promotional announcements.

George Dessart, *Public-Service Announcement*, in 3 ENCYCLOPEDIA OF TELEVISION 1849 (Horace Newcomb ed., 2d ed. 2004). Stations are not required to broadcast PSAs. However, a station's choice to do so provides some evidence of its discharging its "public interest" requirements.

The practice of airing PSAs originated during World War II. Broadcasters and advertising agencies joined to produce the War Ad Council and collaborated with the Office of War Information to distribute homefront propaganda. *See id.* The Ad Council, which dropped "War" from its name after World War II, *id.* at 1849-50, is still a dominant producer of PSAs. *See Marshaling the Ad Industry's Resources*, BROADCASTING & CABLE, Sept. 29, 2003, at 26 (describing the Ad Council as "the nation's biggest producer of public service announcements").

⁴⁷⁷ This is an open scientific question. Of course, it is possible that the approach could backfire or have unintended consequences. The viewing public could see such announcements as political correctness and consequently backlash. Given dissociation, however, it is not obvious that explicit backlash will necessarily prevent disinfection on the implicit level. Also, Emma Jordan has raised to me the danger of producing disinfectant-resistant strains of viruses. Or even more creatively, the body politic — given what it is — may (mis)identify the "disinfectant" as the virus. These are intriguing possibilities, but here, we need to remember that "virus" is a metaphor. I accept that there may be negative consequences to d-PSAs. A full cost-benefit accounting, which would be necessary before any actual policy implementation, must await further scientific experimentation.

⁴⁷⁸ In 2001, adults 18 or older watched 4 hours and 33 minutes of television per day. *See U.S. CENSUS BUREAU, STATISTICAL ABSTRACT OF THE UNITED STATES: 2003*, at 720 tbl.1125, available at <http://www.census.gov/prod/2004pubs/03statab/inforcomm.pdf>.

in court. For example, broadcast stations have never challenged the current children's educational television programming rules, which strongly encourage broadcasters to show three hours of such programming per week.⁴⁷⁹

On the other hand, a licensee may voluntarily broadcast these d-PSAs, as an exercise of its editorial judgment in discharging its "public interest" responsibilities. This describes the current practice of airing PSAs, which broadcasters happily document as billions of dollars worth of airtime donated to the community.⁴⁸⁰ Such voluntary broadcasting would not constitute state action, which means that no constitutional challenge would be available.

Second, consider notice. On the one hand, experiments such as the Computer Crash study suggest that d-PSAs could work even if they are subliminal (compare again with fluoridation or an antivirus software package that automatically updates itself weekly, without user intervention). They would also have the attractive characteristic of not taking up advertising time. Of course, subliminal programming would never be tolerated by the American people.⁴⁸¹

On the other hand, these announcements could be supraliminal, similar to current PSAs. What might be different is that these announcements could last just a few seconds, more like scenes from a fast-cut music video than a lugubrious documentary.⁴⁸² Imagine the attractive Denzel Washington saying, "Be fair. That's what MLK would do." Cut to Tiger Woods echoing the same. Cut to a picture of Martin Luther King, Jr., civil rights marchers, and a United States

⁴⁷⁹ See *In the Matter of Policies and Rules Concerning Children's Television Programming*, 11 F.C.C.R. 10,660, 10,718-21 (1996) (discussing processing guidelines).

⁴⁸⁰ National Association of Broadcasters President Edward Fritts claimed that, in a recent census, broadcasters donated \$9.9 billion worth of airtime for PSAs. See Edward O. Fritts, *TV: A Vast Oasis of Public Interest Programming*, 55 FED. COMM. L.J. 511, 512 (2003). I am skeptical of the figure.

⁴⁸¹ There are no federal or state statutes directly banning subliminal messages or advertisements. See Scot Silvergate, Comment, *Subliminal Perception and the First Amendment: Yelling Fire in a Crowded Mind?*, 44 U. MIAMI L. REV. 1243, 1266 (1990). At the federal level, the FCC has announced publicly that it believes subliminals are against the public interest, but the Commission has never issued regulations or a more formal policy statement. See Harry Schiller, Note, *First Amendment Dialogue and Subliminal Messages*, 11 N.Y.U. REV. L. & SOC. CHANGE 331, 359 (1983). With respect to subliminal advertisements, the Lanham Act's ban on unfair trade practices may apply, although there have been no examples of such litigation. Nicole Grattan Pearson, Note, *Subliminal Speech: Is it Worthy of First Amendment Protection?*, 4 S. CAL. INTERDISC. L.J. 775, 783-84 (1995). Also, it is possible that the FTC could invoke its general enforcement authority against unfair and deceptive practices, and the Bureau of Alcohol, Tobacco, Firearms, and Explosives has issued express regulations against subliminal advertisements for products within its jurisdiction. See Silvergate, *supra*, at 1268-69. Finally, the National Association of Broadcasters and the major television networks are officially on record against subliminals. See *id.* at 1268; Pearson, *supra*, at 783; Schiller, *supra*, at 354.

⁴⁸² See *supra* note 400 (exploring the differences between explicit and implicit engagement with implicit biases).

flag. All in 5 seconds.⁴⁸³ Perhaps just in the lower right corner of the television screen, while the local news credits are rolling. Hundreds of such d-PSAs,⁴⁸⁴ using whatever images the social cognitionists find most effective, could be cycled throughout the day.⁴⁸⁵

Which images would do what work raises intriguing cultural and political questions — indeed, this is where cultural studies meets social cognition. Consider, for example, whether Jesse Jackson would increase or decrease implicit bias for most Americans. What about Al Sharpton? Before the 2004 Democratic Convention? After? What about Malcolm X? Would a message with Thurgood Marshall even be recognized by a majority of high school students? And what images are even available to address anti-Asian nativistic bias? In IATs testing the association of Asian Americans with “American” as opposed to “Foreign,” Asian Americans are consistently associated with “Foreign.” These results reflect the “forever foreign” stereotype that Asian Americans face, which explains why I still not infrequently get complimented on my English. Even when the IAT is run with names and pictures of famous, recognizable, and highly assimilated Asian Americans (such as Kristi Yamaguchi and Connie Chung) against famous White non-Americans (such as German Katarina Witt and Frenchman Gérard Depardieu), Asians lose.⁴⁸⁶ This explains how MSNBC could

⁴⁸³ Shorter is clearly better, from the station’s perspective. The conventional wisdom for PSA placement is that ten-second PSAs “get the most exposure in heavy viewing periods, particularly in major cities.” Margie Goldsmith, PSA Research Center, *How To Get Results with PSAs*, at <http://www.psaresearch.com/bib4208.html> (last visited Feb. 13, 2005).

⁴⁸⁴ By one count, a TV station averages 200 PSAs per week. Jack McGuire, PSA Research Center, *Let’s Clear The Air About Public Service Announcements*, at <http://www.psaresearch.com/cafr1020.html> (last visited Feb. 13, 2005). A study by the Kaiser Family Foundation found:

[B]roadcast and cable television networks donate an average of 15 seconds an hour to air public service ads (PSAs), representing just under one-half of one percent (0.4%) of all airtime. By comparison, 25 percent of airtime is filled with paid advertising and promotions. A significant portion of the donated airtime (43%) is during the late night hours between midnight and 6 a.m.; nine percent is in prime time.

Press Release, Kaiser Family Foundation, *Television Donates an Average of 15 Seconds an Hour to Public Service Advertising; 43% of Donated Time Is in Overnight Hours* (Feb. 21, 2002), available at http://www.kaisernetwork.org/health_cast/uploaded_files/PSA_News_Release.pdf. The practice of embedding network stars in the PSAs could also create win-win opportunities to promote entertainment programming while furthering the public interest. See Editorial, *Viewpoint: TV’s Foolish PSA Wrangle: Value Added*, ADVERTISING AGE, May 19, 1997, at 36 (noting use of TV stars and decrying drop of old-fashioned PSAs “to less than 30 prime-time seconds a night”).

⁴⁸⁵ Yet another variation is a debiasing image appearing in the lower right of the television screen, replacing current station logos. This would also have the benefit of not taking up advertisement time.

⁴⁸⁶ See Devos & Banaji, *supra* note 103, at 24–25 (finding implicit bias even when participants had actual knowledge about who was American and who was foreign).

easily headline Tara Lipinski's defeat of Michelle Kwan, "American Beats Kwan" — when they were both Americans.⁴⁸⁷

Finally, consider how the audience could manifest its assent to receiving these d-PSAs. Suppose we include an implicit bias option in the next generation V-chip, which is embedded in our television sets. Then, only those viewers who consented to d-PSAs would be exposed to them. Those who thought this was mind control could avoid them entirely.⁴⁸⁸ Choice could be exercised through an "opt-out" or an "opt-in" regime. In an "opt-out" regime, if the viewer does nothing, she will be exposed to these announcements; by contrast, in an "opt-in" regime, the viewer must take some affirmative action to program her V-chip to gain access to these announcements.⁴⁸⁹

TABLE I. DISINFECTING PSA OPTIONS

OPTION	STATE ACTION (mandatory/voluntary)	NOTICE (subliminal/supraliminal)	CONSENT (opt-out/opt-in)
1	mandatory	subliminal	opt-out
2	mandatory	subliminal	opt-in
3	mandatory	supraliminal	opt-out
4	mandatory	supraliminal	opt-in
5	voluntary	subliminal	opt-out
6	voluntary	subliminal	opt-in
7	voluntary	supraliminal	opt-out
8	voluntary	supraliminal	opt-in

Table I lists the possible disinfection strategies. They range from most to least disconcerting. Option 1, which is mandatory on the licensee, subliminal, and requires opt-out by the viewer, seems Orwellian — although a truly totalitarian state would not tolerate opt-out, not even by turning off the "two-way screen." Thankfully, one could not imagine such a strategy ever being adopted politically or being tolerated constitutionally in the United States.⁴⁹⁰ In sharp contrast, Option 8, a voluntary decision by licensees to broadcast disinfection,

⁴⁸⁷ See Eric Sorensen, *Asian Groups Attack MSNBC Headline Referring to Kwan*, SEATTLE TIMES, Mar. 3, 1998, at A15.

⁴⁸⁸ One might worry that those who need debiasing the most will be least likely to consent. That may be true. But given the substantial dissociation between explicit and implicit bias, many people could benefit from debiasing and have explicit self-conceptions that would embrace d-PSAs.

⁴⁸⁹ The current V-Chip is opt-out. Unless you program your television otherwise, the default rule is that all programs are shown.

⁴⁹⁰ By way of comparison, one could classify our now-defunct Fairness Doctrine as mandatory on the licensee, supraliminal, and with no opt-out or -in for viewers.

which is supraliminal and thereby provides clear notice, and requires an affirmative opt-in for individual viewers to see d-PSAs in the first place, sounds both politically feasible and constitutional. After all, how different is this from current PSAs against smoking or violence? Compare also the decision to produce and broadcast a *Sesame Street* that features positively valenced characters of all races (and species) enjoying integrated neighborhoods that do not reflect any real city in America. Is this not one of the reasons why we opt in to these programs on behalf of our children?

My goal here is not to analyze each option along the metrics of political feasibility, scientific soundness, and constitutional validity. Rather, I have two more modest goals. The first is simply to point out the feasibility of a disinfection strategy using the same vector that caused infection in the first place. Options 7 and 8 would be the most realistic places to start. The second goal is to suggest how various strategies could be implemented in ways that respect individual choice enough to avoid constitutional problems.

Consider, for example, the state action variable. Between “mandatory” and “voluntary” is regulation via raised eyebrow. An excellent example is the V-chip system, in which the broadcast industry “voluntarily” adopted self-labeling grammars and practices, but only in the shadow of threatened regulatory intervention. An equally voluntary adoption of d-PSAs would not be difficult to engineer. Pressure need not even come solely from the state; consider, for example, the NAACP boycotts of recent, nearly all-White primetime network programming.⁴⁹¹

Consider also the opt-out/opt-in distinction. The transaction costs of flipping out of the default rule determine whether the default setting is what I have called “sticky” or “Teflon.”⁴⁹² Given the fact that many VCRs across the nation blink “12:00 a.m.,” there is reason to think that the default “implicit bias” setting on a V-chip, if not especially offensive, would be sticky. By selecting an opt-out regime, one could thus expose a vast majority of the households to disinfection while emphasizing the formal choice to decline exposure. In addition, old-style PSAs could apply social norm pressure for viewers not to opt out of the d-PSAs. One could imagine an avuncular Bill Cosby saying, “It’s the American way. Don’t be biased. Keep it on.”

3. *The Autonomy Objection.* — For some, everything I have said is deeply disturbing. It is an invitation to state manipulation of its citizenry. It is a disrespectful caricature of the human mind, which is not

⁴⁹¹ See Baynes, *supra* note 434, at 293–95.

⁴⁹² See Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 STAN. L. REV. 1193, 1256–69 (1998) (discussing “sticky” and “Teflon” default rules in the context of cyberspace privacy).

a mere computer vulnerable to viruses. It is a direct affront to the individual's autonomy. This is the "autonomy objection." It is strongly felt. It is understandable. It is untenable.

According to Richard Fallon, *descriptive* autonomy⁴⁹³ is an empirical measure of the degree that an individual enjoys: "(i) critical and self-critical ability, (ii) competence to act, (iii) sufficient options, and (iv) independence of coercion and manipulation."⁴⁹⁴ I pick Fallon's conceptualization not as a strawperson, but as a precise, comprehensible articulation of an otherwise slippery idea; moreover, Fallon's explication pays special attention to the First Amendment context, which is highly relevant to the broadcast application I am making. With Fallon's clarification, the autonomy objection can be stated this way: my call for state intervention in broadcast undermines descriptive autonomy principally by encouraging coercion⁴⁹⁵ and manipulation,⁴⁹⁶ thus undermining the fourth prong of descriptive autonomy.

This objection incorrectly supposes that prior to state intervention (to build a firewall or to broadcast disinfection), we existed in some virginal state without coercion or manipulation. But I have demonstrated that Trojan Horses are being broadcast right now, everywhere, in late-breaking, saturation coverage. The Trojan Horses have been beaming into our brains since we were old enough to be parked in front of a television.⁴⁹⁷ Private actors have always been engaging us, sometimes unknowingly, sometimes shrewdly, on the implicit level. My recommendations are only to counter implicit fire with implicit fire.⁴⁹⁸

⁴⁹³ Richard H. Fallon, Jr., *Two Senses of Autonomy*, 46 STAN. L. REV. 875, 877 (1994) ("[Descriptive autonomy] refers to people's actual condition and signifies the extent to which they are meaningfully 'self-governed' in a universe shaped by causal forces.").

⁴⁹⁴ *Id.* at 886.

⁴⁹⁵ *Id.* at 889 ("[C]oercion' connotes the deliberate and wrongful subjecting of one human being to the will of another or domination that disrespects the other's equal moral worth.").

⁴⁹⁶ *Id.* ("Manipulation' involves a similar judgment of moral wrongness. It implies a perversion of someone's thinking, deciding, or acting." (footnote omitted)).

⁴⁹⁷ See George Gerbner et al., *Growing Up with Television: The Cultivation Perspective*, in MEDIA EFFECTS: ADVANCES IN THEORY AND RESEARCH 17, 17-37 (Jennings Bryant & Dolf Zillman eds., 1994) (describing how television cultivates and socializes all of us, with focus on the authors' "Cultural Indicators" project). See also *id.* at 29-30 (demonstrating that heavy exposure to television increases the tendency to view the real world as violent, mean, and dangerous). Gerbner's research, though, has been controversial. See Barrie Gunter, *The Question of Media Violence*, in MEDIA EFFECTS: ADVANCES IN THEORY AND RESEARCH, *supra*, at 163, 184-86.

⁴⁹⁸ One could respond with a different adage that "two wrongs do not make a right." This objection raises difficult questions about what a truly "neutral" baseline might be. In any event, the strategies I am exploring are designed to "counter" the manipulations of racial meanings that are already taking place in the status quo. I am not simply "piling on" to what is there. Instead, my goal is to try to lessen the power of the racial meanings that are otherwise accumulating. Finally, some worry that using implicit engagements might "backfire" because powerful interests will co-opt these strategies to maintain the status quo. My response is that this objection raises no pref-

Consider it this way. Suppose that a new breed of implicit advertisements, just this side of subliminal, were widely used. Suppose further that they were effective in altering our implicit associations between, say, ourselves and a particular commodity, aesthetic, or brand. After all, one could joke, what else would make us pay hundreds of dollars for ratty jeans that have dirt rubbed into the thighs? How would we respond if the government outlawed these implicit advertisements on the airwaves? Would we suggest that there was no manipulation before, but suddenly with the government ban, we were being manipulated? If the prohibition were unconstitutional, what if the government responded with supraliminal warning messages in the lower right corner of the television screen that shielded viewers from the “viral marketing” payload? Again, would we suddenly protest this warning as mind control when before we were cocksure of our autonomy? The option of no manipulation — given the best science known to us — is not an option. The status quo is hardly “neutral and natural.”⁴⁹⁹ We must decide which threatens autonomy more, Trojan Horses or their remedy.

The predictable response is to retreat to the public/private distinction, central to our constitutional jurisprudence. Privately launched Trojan Horses are tolerable; public remediation, by contrast, is not. But from an autonomy perspective, what is the empirically grounded justification for maintaining this sharp and inflexible distinction? And remember how we got onto the subject of local news in the first place: the FCC, implementing Congress’s “public interest” standard, recently altered market structure to increase the production of local news. State action that *increases* virus production is already manifest. On one level, then, my call is simply for the FCC to do no harm.

State action is manifest in other ways — for example, in protecting the quasi-property rights of licensees.⁵⁰⁰ If I broadcast debiasing counterspeech, in exercise of my freedom of expression, on the local ABC affiliate’s frequency, I would be criminally prosecuted *by the state* for pirate radio. Finally, if form remains supremely important, I have delineated various state strategies that could be made to look sufficiently private and voluntary so as not to trip the state action wire.⁵⁰¹

My response to the autonomy objection strengthens as we consider the other prongs of autonomy Fallon identifies. Consider, for example,

erable alternative, because it is naive to think that by unilaterally renouncing any consideration of implicit cognitive processes, those on the “other side,” however defined, will follow suit.

⁴⁹⁹ See Fallon, *supra* note 493, at 883 (footnote omitted) (describing the mistake of “negative libertarians” espousing a descriptive autonomy conception of the First Amendment).

⁵⁰⁰ See generally JOSEPH WILLIAM SINGER, ENTITLEMENT: THE PARADOXES OF PROPERTY 32 (2000) (pointing out that regulation is necessary for property rights).

⁵⁰¹ See *supra* pp. 1581, 1583.

the third prong of “sufficient options.” As Fallon concedes, it is hard to know how much richness and variation is required for autonomy, but the experience of options should not be a purely subjective inquiry.⁵⁰² In other words, as Fallon puts it, the “happy slave” is not autonomous.⁵⁰³ This, of course, raises the question of preference formation. And there is now increasing evidence that implicit biases influence what we end up “preferring.” In other words, general background stereotypes can create certain linkages (for example, Professor = White) that may deter individuals from preferring particular roles, ideas, or futures.⁵⁰⁴ By responding to the implicit bias findings, we can contest this invisible narrowing of options and imagined futures.

Consider, finally, how an honest engagement with the implicit bias literature promotes the first two prongs of Fallon’s autonomy test: “critical and self-critical ability” and “competence to act.” As Fallon explains, the first prong goes beyond mere instrumental, means-ends rationality. Rather, it requires some self-critical evaluation of the “tastes, goals, and desires” we have at any given moment.⁵⁰⁵ The second prong requires the ability to translate that self-critical deliberation into action in ways that, for example, an addict trying to overcome her addiction cannot.⁵⁰⁶ Now focus on one of the central findings of social cognition: *dissociation*. Even when we have explicitly decided in favor of racial equality, something implicit within us does not follow. It is as if the racial meanings ubiquitous in the culture and embedded in our histories exert a deterministic drag that prevents us from fully adopting self-defining choices about fairness and basic human decency. Moreover, these implicit biases, like viruses, undermine our competency to act in accordance with our self-reflectively chosen ends. We do not want to treat minorities poorly, but sometimes we cannot help doing so. Viewed in this light, social strategies that decrease the dissociation that we as a society collectively experience should be seen as autonomy-reinforcing. This tactic parallels the attempt to create alignment between second-order desires (desires about desires) with first-order desires, which on one philosophical account is the essence of

⁵⁰² See Fallon, *supra* note 493, at 888–89.

⁵⁰³ *Id.* at 888.

⁵⁰⁴ See generally Brian A. Nosek et al., *Math = Male, Me = Female, Therefore Math ≠ Me*, 83 J. PERSONALITY & SOC. PSYCHOL. 44 (2002) (demonstrating how background cultural stereotypes that math is not a female subject can discourage women from wanting to study math in the first place).

⁵⁰⁵ See Fallon, *supra* note 493, at 887–88.

⁵⁰⁶ See *id.* at 888.

freedom of will.⁵⁰⁷ How could it be against autonomy to bring our implicit thoughts in line with our explicit ones?⁵⁰⁸

Perhaps our ambivalence remains because the autonomy objection is really instead about *ascriptive* autonomy. Ascriptive autonomy is a metaphysical foundation ascribed to all competent adults⁵⁰⁹ that warrants personal sovereignty, at least over all self-regarding decisions.⁵¹⁰ In contrast to the descriptive variant, ascriptive autonomy is not so much a situated, empirical measure; rather, it is an article of faith, often instantiated in law, consistent with our moral intuitions of human dignity and agency, and fundamentally resentful of paternalism of any sort. Fallon concludes that ascriptive autonomy should be seen as a “fundamental moral and political value” of which our constitutional law, including First Amendment theory, should take account.⁵¹¹

But, in the end, it is unclear how an ascriptive autonomy objection might go. According to Fallon’s sensible formulation, its bite is limited to those realms that are at least arguably self-regarding.⁵¹² But Part I demonstrates that implicit biases have behavioral consequences that are by no means limited to one’s self. Remember Amadou Diallo. If the ascriptive autonomy objection is extended beyond this self-regarding realm, I do not understand the normative grounds on which an unshakable “faith” in autonomy should be allowed to perpetuate an unjust status quo, when the best self-understanding we have suggests that we could do better. Faith in ascriptive autonomy, that there are no ghosts in the machine, undoubtedly produces a useful and adaptive

⁵⁰⁷ See Harry G. Frankfurt, *Freedom of the Will and the Concept of a Person*, 68 J. PHIL. 5, 14–17 (1971).

⁵⁰⁸ It is possible that for some individuals, explicit bias against a social category may be quite high, whereas implicit bias may be zero. According to the preceding account of autonomy, it would be autonomy-reinforcing to bring her implicit bias in line with her explicit bias, and in autonomy terms, that process could involve either moving the implicit bias toward her explicit bias (making implicit bias more negative) or vice versa (making explicit bias go to zero). As a matter of social policy, we would prefer the latter. But I recognize that from a purely autonomy-reinforcing position, one might reasonably be indifferent to the two directions of alignment.

⁵⁰⁹ One reason why it has been easier to implement content control strategies framed as pro-children policies, such as suppression of indecency and violence, is that in our legal culture, ascriptive autonomy does not apply to minors.

⁵¹⁰ See Fallon, *supra* note 493, at 878 (“[Ascriptive] autonomy represents the purported metaphysical foundation of people’s capacity and also their right to make and act on their own decisions, even if those decisions are ill-considered or substantively unwise.”).

⁵¹¹ *Id.* at 893.

⁵¹² See *id.* at 893–94.

“anxiety buffer.”⁵¹³ But the myth of autonomy should not be used to deny the reality of racial bias.⁵¹⁴

CONCLUSION

Indulge me in some science fiction.

* * * * *

It is the year 2200. Extraordinary advances in neural network computing and nanotechnologies have allowed us to implant “augmented intelligence,” called <augI>, into our brains. IntelliDyne Corp. has patented this and related technologies. For example, standard retina upgrades allow us to see ultraviolet and infrared rays, zoom into fine detail, and “see” radio frequency identification (RFID) signals that radiate from nearly all objects. These data are recorded 24/7 in digital memory, with processors engaged in real-time pattern recognition and self-initiated, intelligent queries to massively interlinked, ubiquitously available databases. All this in a carbon-wrapped piece of silicon circuitry weighing three grams and implanted in the limbic system, next to the amygdala. These days, no one forgets anyone’s name, face, marital status, or even sexual orientation; <augI> presents that information to us inside of 200 milliseconds.⁵¹⁵

IntelliDyne’s creative breakthrough was in user interface. It recognized that users would quickly be overwhelmed by information if it were provided explicitly, in semantic form. Therefore, calculation results had to be provided implicitly, through sight, sound, smell, touch, taste, as well as mood. For example, first responders equipped with public safety <augI> can actually smell radiation. Law enforcement <augI> not only enables officers to run identity checks from a distance, but also biochemically quickens physiological responses whenever officers are in “threat mode.” When <augI> senses danger, for example by recognizing a face and associating it with an outstanding felony warrant, officers see increasingly deeper hues of red as a function of estimated threat and they experience rushes of adrenaline.

IntelliDyne’s profits are immense. Demand for the academic <augI> is insatiable as parents of all economic classes struggle to implant their children before high-stakes testing. Managers have become accustomed to using <augI> to assist financial, employment, and strategic decisionmaking. Productivity is bounding for those nation-states

⁵¹³ See Donald C. Langevoort, *Taking Myths Seriously: An Essay for Lawyers*, 74 CHI.-KENT L. REV. 1569, 1571–77 (2000) (describing cognitive biases that help us increase certainty and confidence, and thereby decrease anxiety).

⁵¹⁴ Cf. Langevoort, *supra* note 21, at 1526 (suggesting that the rationality assumption plays a mythical, soothing role).

⁵¹⁵ Cf. Jerry Kang & Dana Cuff, *Pervasive Computing: Embedding the Public Sphere*, 62 WASH. & LEE L. REV. (forthcoming Mar. 2005) (describing elements of pervasive computing).

that have widely adopted these technologies. Society seems to be on the edge of a new cyber-assisted, rational, hyperefficient utopia.

Except for rumors that simply refuse to die. In the blogosphere, many rant that there are bugs in the IntelliDyne system. Some worry whether <augI> has been hacked and whether Trojan Horse viruses have unexpectedly entered our brains. Critics say that <augI> has a systematic bias, that it makes reproducible errors. The rumors are growing, especially among the Browns. Just last month, law enforcement shot another Brown who was not armed and lacked a criminal record. Digital audit trails revealed that the officers were operating under extreme threat mode.

After receiving special congressional authorization, university scientists are permitted to decrypt, reverse engineer, then examine IntelliDyne's <augI> neural network modules. Initial research suggests that these modules produce different results among what should be arbitrary differences in visual stimulus. For example, in a stunning finding, scientists revealed that "threat mode" is quicker to identify guns when they are held by Browns than when they are held by Whites. Under the highest "threat" level, these modules make it easier to shoot Browns than to shoot Whites. Scientists repeatedly confirm these and other troubling results. For instance, the academic <augI> modules seem to malfunction under particular circumstances, especially when implanted in Browns.

When Congress subpoenas the IntelliDyne engineers for an inquiry, they explain that neural networks are not "programmed" to execute specific lines of code that instruct mechanically what should be done if certain conditions are satisfied. Rather, they are "programmed" through millions of exposures to exemplar cases, which over countless iterations set the appropriate strength of linkages between neural network nodes. The engineers are confident that the programming exposures are appropriate and representative.

There is growing pressure to stop implanting <augI> modules or to cure them somehow of these biases. At the same time, many strongly resist, saying that the evidence of bias is equivocal at best and that conspiracy theorists should not be heeded. Anyway, as IntelliDyne repeats, <augI> merely "assists" decisions that are made under the conscious control of human beings exercising independent, fully human judgment. As their advertisements go, "Assisted Autonomy: It's Your Choice." The most vociferous supporters assert that it is their fundamental human right to implant <augI> if they so choose, and Big Brother must not interfere with private choices. A catastrophic show-down looms.

* * * * *

It is only the year 2005, and thankfully there is no IntelliDyne and there is no <augI>. But the social cognition research I have introduced demonstrates that we may not need anything foreign stuck in

our brains for disturbing biases to be present. If man-made neural implants were demonstrably faulty because of bad programming or virus infection, would we not aggressively demand governmental intervention? Regulatory approvals would be revoked; law suits would be filed; legislation would be passed. But of course, we are talking instead about our brains and not some external implant, which changes our response, perhaps radically. Maybe this is entirely appropriate. I am asking why.

The arc of this Article has been long, and given its multiple goals, it has been more evocative than comprehensive. A primary goal is to make the case for using social cognition in critical race studies. In the 1980s and 1990s, debates raged about the best or most appropriate methodology with which to engage in “criticism” of law and legal institutions on matters of racial equality. Countless articles explored, for example, whether narrative defended through postmodernism would be the best or only way.⁵¹⁶ Countless articles explored whether minority scholars did or should have preferred standing to make these inquiries. We have learned from those debates, and the time has come to move on and add things new. This Article has been an attempt to demonstrate how and why that should be done.

The benefits will not flow only in one direction, from science into law. Instead, legal analysts who are subject to different craft norms can apply and extend the science into the policy realm in ways that social cognitionists cannot. Less instrumentally, as outsiders, we can identify scientific blindspots. For example, the lack of exchange between psychologists who study stereotype threat and those who study implicit bias is perplexing when the same or related cognitive processes are likely to be at the heart of both phenomena. The lack of exchange between political scientists and social cognitionists is also odd; for instance, why do political scientists still measure only explicit bias after exposing participants to differently edited news stories?⁵¹⁷ Adding one round of the IAT would be enormously illuminating. The upshot is a call for a new school of thought called “behavioral realism,”⁵¹⁸ in which legal analysts, social cognitionists (with emphases in implicit bias and stereotype threat literatures), evolutionary psychologists, neurobiologists, computer scientists, political scientists, and behavioral

⁵¹⁶ This is still a live debate. See, e.g., Kevin R. Johnson, Commentary, *Roll Over Beethoven: “A Critical Examination of Recent Writing About Race”*, 82 TEX. L. REV. 717 (2004) (commenting on Richard Delgado’s negative review of a critical race theory book that emphasizes idealist discourse over materialist analysis).

⁵¹⁷ There is nothing especially wrong with using explicit bias measures, especially when political scientists discover interesting correlations between these measures and the direction and magnitude of participants’ responses. Still, a rich underlayer of analysis and consequence is missing when implicit cognitive processes are not measured.

⁵¹⁸ See *supra* note 21.

(law and) economists cooperate to deepen our understanding of human behavior generally and racial mechanics specifically, with an eye toward practical solutions.⁵¹⁹ The next generation of critical race scholars should be at the forefront of this endeavor and not in some rear-guard action. Sitting on the back of this bus is not an option.

A more modest goal of this Article is to bridge divides within the law itself. As in *Cyber-race*,⁵²⁰ I am trying to cajole legal scholars working in cyberlaw and communications to engage with race as well as other social categories of subordination. At the same time, I am trying to persuade race scholars to select unconventional points of entry by adopting unorthodox subjects, metaphors, and analytic tools. The cross-fertilization should help us think things anew. The crucible for this Article has been the FCC's recent mass media ownership deregulation — specifically the Commission's fixation on local news. Local news explicitly furthers the public interest, but its fetish for violent crime makes it a Trojan Horse, a “thing that undermines from within.”⁵²¹

I have made a solid case for recoding the FCC's definition of the public interest to decrease its reliance on local news. I recognize that counting hours of local news is simple, but something can be both simple and wrong. My ideas about capping crime stories and broadcasting d-PSAs are purposefully labeled “thought experiments.” The value of these inquiries is that they prepare us for actual conversations that will come soon. Moreover, they hint at another sort of implicit bias running beneath the surface of our law. The difference between our complacency about modulating indecency and our anxiety about modulating bias speaks volumes about our culture, politics, and law. Not everything it says is about our explicit commitments. Rather, much of our doctrine might be best explained by our implicit beliefs, which have remained shielded from critical examination.

I close with a caution and a call. The caution is that the remarkable science of implicit bias could draw all of our interest and attention. But implicit bias is not the only source of pervasive and persistent inequalities among social groups. Explicit bias still thrives in

⁵¹⁹ For a thoughtful commentary on the limitations of behavioral science, including behavioral law and economics and social norm theories, see Rostain, *supra* note 255.

⁵²⁰ See Kang, *supra* note 19, at 1207; see also Kang, *supra* note 36 (exploring the relationship between e-voting and race).

⁵²¹ 18 THE OXFORD ENGLISH DICTIONARY 574 (2d ed. 1989) (“*Trojan horse*: according to epic tradition, the hollow wooden horse in which Greeks were concealed to enter Troy; *fig.* a person, device, etc., insinuated to bring about an enemy's downfall; a person or thing that undermines from within . . .”).

many circles.⁵²² Durable inequality⁵²³ may also be maintained by structural arrangements that are no longer tightly connected to bias, implicit or explicit.⁵²⁴ Implicit bias should not circumscribe the content of our concerns.

Mahzarin Banaji, a leading scientist in the field of implicit bias, has suggested that “one measure of the evolution of a society may indeed be the degree of separation between conscious and unconscious attitudes — that is, the degree to which primitive implicit evaluations that disfavor certain social groups or outgroups are explicitly corrected at the conscious level at which control is possible.”⁵²⁵ Although my response to the autonomy objection was framed at the individual level, Banaji’s insight restates that response at the level of entire societies. Maybe this alignment between the explicit and implicit cannot be reached, at least not perfectly. Evolutionary psychology will surely have its say. Still, achieving this convergence is our challenge. It is our call.

⁵²² Consider, for example, the recent survey of 135 undergraduates from Yale — hardly a site of White supremacy. As Devos and Banaji report, even after being explicitly told to “consider individuals from each ethnic group who were born in the U.S., lived in the U.S., and were U.S. citizens,” they rated Asian Americans as much less “American” than White and Black Americans. See Devos & Banaji, *supra* note 103 (manuscript at 10) (reporting, on a seven-point scale, that Whites were American with a mean of 6.53, African Americans were American with a mean of 6.26, but Asian Americans were American with a mean of only 5.49; $p < 0.001$).

⁵²³ See CHARLES TILLY, *DURABLE INEQUALITY* 6 (1998) (discussing “the causes, uses, structures, and effects of categorical inequality”).

⁵²⁴ Legal doctrines, such as certain strands of disparate impact theory, have attempted to respond to such arrangements.

⁵²⁵ Banaji, *supra* note 108, at 134.