

# California Law Review

---

VOL. 83

MAY 1995

No. 3

---

Copyright © 1995 by California Law Review, Inc.

## Stereotypes and Prejudice: Helping Legal Decisionmakers Break the Prejudice Habit

Jody Armour†

*Courts are reluctant to allow attorneys to bring the issue of prejudice into the open at trial. Judges have barred attorneys who represent socially marginalized clients from pointing out that their client comes from a stereotyped group on the grounds that such references "play to the prejudices of the jury." Although many legal commentators stress the need to reduce bias in legal proceedings, few commentators have considered whether the colorblind formalism advocated by courts actually helps legal decision makers avoid discrimination. This Article argues that colorblind formalism is counterproductive in reducing discrimination, and it develops an empirically grounded framework for combating unconscious discrimination. After demonstrating the shortcomings of the "unconscious racism" model that has been the focus of much debate among legal commentators, the author draws on recent research in social and cognitive psychology to distinguish between stereotypes and prejudice. The Article defines stereotypes as well-learned internal associations about social groups that are governed by automatic cognitive processes. In contrast, prejudice consists of a set of*

---

Copyright © 1995 California Law Review, Inc.

† Associate Professor, University of Pittsburgh School of Law. A.B. Harvard University, 1982; J.D. University of California at Berkeley, 1986. This Article is dedicated to my grandmother, Addie Marion Armour, a proud descendant of the African diaspora who never lets us forget our true and ancient properties. I owe special thanks to Frank McClelland, Martha Chamallas, Veronica Hobbs, Taunya Banks, Jules Lobel, Welsh White, Rhonda Wasserman, Bernard Hibbitts and to the pioneers of Critical Race Scholarship, whose dogged testing of our tacit assumptions made room in the academy for the kind of research embodied in this article. Thanks also to Ruth Colker and Robert Harper for helpful comments, to Rhonda Evans, Dominick Lee, Shakura Sabur, and Bomi Anise for their research assistance, and to LuAnn Driscoll, Karen Knochel, Darleen Mocello, Carolyn Rohan and Barbara Salopek for word processing assistance. This project was supported by an Alumni Summer Research Stipend from the University of Pittsburgh School of Law.

*conscious personal beliefs. The author suggests that automatic negative responses to stereotyped groups can be controlled by activating controlled cognitive processes. Thus, through careful attention and conscious effort, a non-prejudiced person can suppress an ingrained stereotype. As a result, the author concludes that references in court that challenge jurors to reexamine and resist their automatic discriminatory tendencies may actually enhance, rather than impede, the rationality and fairness of legal proceedings.*

#### INTRODUCTION

A certain brand of judicial formalism promotes the very discrimination it purports to eliminate. Such self-defeating formalism typically takes the following form: a court forbids references during formal legal proceedings to a social characteristic such as race, gender, or sexual orientation, because that trait historically has been a basis for invidious discrimination.<sup>1</sup> As a result, legal decisionmakers are less able to control their discriminatory responses than they would be if they confronted and dealt with their antagonism towards the characteristic in question. A dramatic cinematic depiction of this type of judicial formalism appears in several courtroom scenes in the movie *Philadelphia*. The film—inspired by a true story<sup>2</sup>—concerns a successful gay attorney, Andrew Beckett, who is wrongfully discharged by his law firm because of his sexual orientation. In the following scene, Beckett's attorney is cross-examining a firm employee who was involved in the conspiracy to discharge Beckett wrongfully:

PLAINT. ATTY. Are you a homosexual?

WITNESS What?

PLAINT. ATTY. Are you a homosexual?

Answer the question.

Are you a homo? Are you a faggot? . . .

fairy . . . booty snatcher . . . rump-roaster. Are you gay?

DEF. ATTY. Where did this come from?

[The witness's] sexual orientation has nothing to do with this case.

PLAINT. ATTY. Your honor, everybody in this courtroom is thinking about sexual orientation, you know, sexual preference, whatever you want to call it. Who does what to whom, and how they do it. I mean, they're looking at Andrew Beckett [plaint.], they're thinking about it. They're

1. See, e.g., *infra* notes 168-171 and accompanying text.

2. *Cavagnuolo v. Baker & McKenzie*, No. 1B-E-D-86-115824 (N.Y. State Div. Human Rights Dec. 17, 1993).

looking at Mr. Wheeler [senior partner], Ms. Cornini [defense counsel], even you your honor. They're wondering about it. I mean, hey, trust me, I know that they are looking at me and thinking about it. So let's just get it out in the open, let's, let's get it out of the closet, because this case is not just about AIDS is it? So let's talk about what this case is really all about, the general public's hatred, our loathing, our fear of homosexuals, and how that climate of hatred and fear translated into the firing of this particular homosexual, my client, Andrew Beckett.

JUDGE

In this courtroom, justice is blind to matters of race, creed, color, religion, and sexual orientation.

PLAINT. ATTY. With all due respect your honor, we don't live in this courtroom though, do we?

JUDGE

No, we don't. However, as regards this witness, I'm going to sustain the defense's objection.<sup>3</sup>

Unfortunately, the formalism exemplified by this fictional judge's reaction to references to sexual orientation reflects real life judicial resistance to attorneys' attempts to bring the issue of prejudice into the open at trial. For example, in *Jackson v. Chicago Transit Authority*,<sup>4</sup> a black plaintiff brought a negligence action against a municipal corporation for personal injuries sustained when the bus he boarded collided with a truck. During his closing argument, the plaintiff's counsel "alluded to the fact that his client was Negro, as contrasted to the jurors, the attorneys and the court itself, who were all Caucasians."<sup>5</sup> The jury returned a verdict for the plaintiff, but the appellate court granted the defendants a new trial on the ground that such a racial reference "should not be made before any tribunal. *It is an unmitigated appeal to prejudice and its effect could only be destructive of the proper administration of justice.*"<sup>6</sup>

In characterizing the reference by the plaintiff's counsel to his client's racial identity as a case of playing to the prejudices of the jury, the *Jackson* court ignores a critical distinction between racial references that subvert the rationality of the fact-finding process and racial references that actually enhance the rationality and fairness of the fact-finding process. Worse still,

---

3. PHILADELPHIA (Tristar Pictures, Inc. 1993).

4. 273 N.E.2d 748 (Ill. App. Ct. 1971).

5. *Id.* at 751.

6. *Id.* (emphasis added).

this court's superficial analysis has gained legitimacy and wide currency through its use and endorsement by the authors of *The Torts Process*,<sup>7</sup> a popular torts casebook. Citing *Jackson* as authority, the casebook authors, James Henderson, Richard Pearson and John Siliciano, assert that "a lawyer's effort to invoke jury sympathy for a client based on such characteristics (as race, nationality, or ethnic background) is as objectionable as an appeal to be unsympathetic to the other side for the same reason."<sup>8</sup> Henderson, Pearson, and Siliciano point to several professional responsibility standards that are relevant to determining whether the attorney in *Jackson* acted appropriately in adverting to his client's racial identity.<sup>9</sup>

Despite Henderson, Pearson, Siliciano, and *Jackson*, however, attorneys frequently challenge factfinders explicitly to resist succumbing to bias in making judgments about members of stereotyped groups. For example, in the recent World Trade Center bombing case, defense attorney Austin Campriello asked the jury to avoid associating stereotypes of Arab and Muslim violence and terrorism with his client.<sup>10</sup> And in a recent capital murder trial, defense attorney Paul Nugent urged the jury not to allow homophobia to distort their deliberations about his client's guilt or innocence.<sup>11</sup> Thus we have a legal process issue that is neither pro-prosecution nor pro-defense: Do arguments based on race, sexual preference, or any other characteristic widely used to stereotype individuals, necessarily

---

7. JAMES HENDERSON ET AL., *THE TORTS PROCESS* (1994).

8. *Id.* at 75. Since the political slant of casebooks can have a lasting impact on the thinking of students and practitioners, I shall develop this point further. I teach torts to first year students from *The Torts Process*, a popular problem-oriented casebook. The cases and materials in the book are organized around fabricated fact-patterns that afford students the opportunity to act out the roles of attorneys. One especially provocative problem involves a storekeeper who claims the privilege of self-defense after shooting a black customer who the storekeeper mistakes for a robber. The storekeeper's privilege turns on whether a reasonable person would have believed he was under attack and the circumstances of the hypothetical shooting are sufficiently ambiguous to provide both sides of the dispute with plausible arguments. I assign students to represent both sides of the dispute and have them present their arguments as closing arguments to a mock jury.

Since studies show that most Americans consider race in assessing the risk of violence a person poses, see, e.g., TOM W. SMITH, *ETHNIC IMAGES*, 4, 9-10, 16 (General Social Survey Topical Report No. 19, 1990) (on file with the author), the role that the black victim's racial identity played in the storekeeper's decision to shoot cries out for recognition. Yet the students assigned to represent the black victim's interest almost always eschew mentioning the racial factor. When I ask why they did not address the issue of the victim's racial identity, they invariably point to suggestions by the authors of the casebook that mentioning the racial factor would both violate the law and constitute a breach of professional responsibility.

9. HENDERSON ET AL., *supra* note 7, at 76. The authors specifically point to the American Bar Association/Bureau of National Affairs' *LAWYER'S MANUAL ON PROFESSIONAL CONDUCT* Rule EC 7-25 ("[A] lawyer should not by subterfuge put before a jury matters which it cannot properly consider."), *id.*, and the American Bar Association's *STANDARDS FOR CRIMINAL JUSTICE*, which contains provisions forbidding prosecutors and defense counsel from using "arguments calculated to inflame the passions or prejudices of the jury." *Id.* (quoting Standards 3-5.8(c), 4-7.8(c) (2d ed., 1986 Supp.)).

10. Patricia Cohen, *WTC Lawyer Cites Bias Against Arabs*, *NEWSDAY*, Feb. 19, 1994, at 18.

11. Todd J. Gillman, *Cook Jury Begins Deliberations, Claims Deadlock After Five Hours*, *THE DALLAS MORNING NEWS*, Feb. 18, 1994, at 32A.

"appeal to prejudice," or instead can some such arguments actually promote the rationality and fairness of the fact-finding process?

I argue that in many situations it may enhance the rationality of the decisionmaking process for attorneys explicitly to challenge factfinders to confront their biases against blacks and members of other stereotyped groups.<sup>12</sup> In arguing that bias reduction among legal decisionmakers is feasible, I must frame a model of prejudice that goes beyond the "unconscious racism" model that informs much of the current legal analyses of prejudice.<sup>13</sup> For although the unconscious racism model certainly does not endorse the kind of colorblind formalism espoused in *Jackson*, neither does it provide theoretical leverage for developing legal strategies aimed at helping decisionmakers to resist unconscious discrimination. Perhaps for this reason, commentators working within the unconscious racism framework have not developed proactive strategies for promoting bias reduction among white legal decisionmakers, but instead have focused on after-the-fact constitutional review of decisions by whites<sup>14</sup> or on techniques to increase the number of black decisionmakers.<sup>15</sup> Race shield laws (modeled after rape shield laws) have also been proposed to prevent the exploitation of racial imagery in criminal trials.<sup>16</sup> Such proposals, however, still presuppose high baseline levels of prejudice in white jurors that racial imagery serves to exacerbate. There is little discussion of techniques for lowering the high baseline level of anti-black bias itself.

I contend that current legal commentators have not explored bias reduction strategies because they have not explored the utility of consciously endorsed, nonprejudiced beliefs in the fight against biased judgments by legal decisionmakers. Current discussions of prejudice in the legal literature view people who report nonprejudiced personal beliefs as either hypocritical or self-deluded. The model of prejudice I will frame, however, posits that many people who report nonprejudiced personal beliefs actually do hold such beliefs. For many, these nonprejudiced per-

---

12. I will focus in particular on racial references in this article because of the uniquely tragic role that race has played in American history and American jurisprudence, and the particularly pressing responsibility of the courts to develop strategies and techniques for freeing legal decisionmakers from racially biased responses.

13. See generally Peggy C. Davis, *Law As Microaggression*, 98 YALE L.J. 1559 (1989); Sheri Lynn Johnson, *Unconscious Racism and the Criminal Law*, 73 CORNELL L. REV. 1016 (1988); Charles R. Lawrence III, *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317 (1987).

14. E.g., Lawrence, *supra* note 13 (criticizing the doctrine of discriminatory purpose established in *Washington v. Davis*, 426 U.S. 299 (1976)).

15. E.g., Sheri Lynn Johnson, *Black Innocence and the White Jury*, 83 MICH. L. REV. 1611, 1691-1708 (1985) (arguing that defendants should have a right to racially similar jurors).

16. E.g., Sheri L. Johnson, *Racial Imagery In Criminal Cases*, 67 TUL. L. REV. 1739, 1797-1803 (1993) (proposing a new "racial imagery shield law"); see also Jody D. Armour, *Race Ipsa Loquitur: Of Reasonable Racists, Intelligent Bayesians, and Involuntary Negrophobes*, 46 STAN. L. REV. 781 (1994) (arguing that race-based evidence of reasonableness should be excluded from cases involving claims of self-defense, even though the racial factor may be formally relevant to such claims).

sonal standards are well-internalized, and techniques that encourage people to activate their nonprejudiced beliefs are the key to reducing discriminatory social judgments about blacks and members of other stereotyped groups. Although I do not dispute the contention of many legal commentators that unconscious discrimination routinely distorts the judgments of legal decisionmakers (to the contrary, I discuss compelling empirical proof that unconscious discrimination frequently biases the social judgments of all Americans), I develop an empirically grounded framework for fashioning techniques that can help legal decisionmakers combat even their unconscious discriminatory tendencies. This framework focuses on both the conscious and unconscious responses that Americans have to persons from stereotyped groups, and to the interplay between these two distinct kinds of responses. Specifically, I suggest ways of activating nonprejudiced beliefs in jury members to counteract their unconscious bias.

In Part I, I analyze the claim by legal commentators that people who profess nonprejudiced personal beliefs are either hypocritical racists or unconscious racists. I argue that these rather pessimistic interpretations of contemporary racial attitudes result from a failure of commentators carefully to distinguish between two distinct sources of responses to blacks and other marginalized social groups—namely, stereotypes and personal beliefs. Conceptualizing stereotypes and personal beliefs as distinct sources of responses to blacks differs from the long-standing tradition of essentially equating the two; that is, a stereotype is often defined as “a set of beliefs about the personal attributes of a group of people.”<sup>17</sup> Yet, drawing upon recent developments in social cognition, which adopts an information-processing model as a means of understanding social judgments, feelings, and behaviors, I argue that the distinction is important because the activation of stereotypes and personal beliefs is governed by different cognitive processes: automatic and controlled processes, respectively. Controlled processes are the key to escaping unconscious discrimination.

In Part II, I elaborate a new model of unconscious discrimination rooted in compelling empirical research. This discussion considers the close relationship between stereotypes and automatic processes and shows how negative responses to members of stereotyped groups are essentially “bad habits,” at least for people who hold nonprejudiced personal beliefs. In Part III, I consider the potential for legal decisionmakers to control (at least temporarily) the discrimination habit by activating their nonprejudiced—or at least egalitarian—personal beliefs. I conclude that it is crucial for courts to distinguish between rationality-enhancing and rationality-subverting uses of racial references to prevent colorblind formalism from

---

17. Wolfgang Stroebe & Chester A. Insko, *Stereotype, Prejudice, and Discrimination: Changing Conceptions in Theory and Research*, in *STEREOTYPING AND PREJUDICE: CHANGING CONCEPTIONS* 3, 5 (Daniel Bar-Tal et al. eds., 1989).

promoting discrimination by factfinders against litigants from stereotyped groups.

# I

## HYPOCRITICAL RACISTS AND AVERSIVE RACISTS

Ethnic attitudes are a part of the social heritage of the developing child. They are transmitted across generations as a component of the accumulated knowledge of society. No person can grow up in a society without learning the prevailing attitudes concerning the major ethnic groups. *In fact, given the polarization of ethnic attitudes, we ought to consider the question of how some people escape being prejudiced.*<sup>18</sup>

This suggestion that some people in our society escape prejudice conflicts with many current discussions of prejudice in the legal literature. Many legal commentators view prejudice as an inevitable outgrowth of our cultural belief system. For example, in an often quoted passage from his seminal article, *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, Professor Charles Lawrence III states:

Americans share a common historical and cultural heritage in which racism has played and still plays a dominant role. Because of this shared experience, we also inevitably share many ideas, attitudes, and beliefs that attach significance to an individual's race and induce negative feelings and opinions about nonwhites. *To the extent that this cultural belief system has influenced all of us, we are all racists.*<sup>19</sup>

This view of prejudice as inevitable and ubiquitous, however, does not square with the survey literature on racial attitudes indicating that prejudice has been declining steadily over the past 40 years. Citing a series of reports on attitudes of white Americans toward black Americans appearing in *Scientific American* between 1956 and 1978, Charles E. Case and Andrew M. Greeley conclude that "there has been a continuous increase in the percent of whites who favor equal treatment for blacks in all areas of American society" since 1942.<sup>20</sup> Furthermore, responding to commentaries suggesting a resurgence of racism in the late 1980s, Charlotte Steeh and Howard Schuman reviewed surveys conducted between 1984 and 1990 on young white adults and concluded that the survey data show that there is "no widespread, systematic decline in liberal racial attitudes among those people entering adulthood from 1960 to 1990."<sup>21</sup>

18. HOWARD J. EHRLICH, *THE SOCIAL PSYCHOLOGY OF PREJUDICE* 110 (1973) (emphasis added).

19. Lawrence, *supra* note 13, at 322 (emphasis added) (footnote omitted).

20. Charles E. Case & Andrew M. Greeley, *Attitudes Toward Racial Equality*, 16 HUMBOLDT J. SOC. REL. 67, 68 (1990).

21. Charlotte Steeh & Howard Schuman, *Young White Adults: Did Racial Attitudes Change in the 1980s?*, 98 AM. J. SOC. 340, 361 (1992).

Legal commentators have reconciled reports indicating diminishing prejudice with the claim that prejudice remains ubiquitous by attacking the validity of the self-evaluations that form the basis of these reports. Specifically, commentators in effect characterize persons who report non-prejudiced personal beliefs as either hypocritical, subrosa racists or unconscious, aversive racists. I will first consider the hypocrisy reading of the survey literature and argue that this rather pessimistic interpretation of the surveys results from a failure to carefully distinguish between stereotypes and prejudice. The distinction between responses based on stereotypes and responses based on prejudiced personal beliefs provides a crucial theoretical tool for understanding how jurors' discriminatory responses to blacks and members of other marginalized groups may be overcome. I will then turn to the aversive racism interpretation of the surveys, examine empirical and conceptual shortcomings of this model of prejudice, and suggest how the dissociation of stereotypes and prejudice helps in the resolution of these shortcomings and in the development of prejudice reduction strategies.

#### A. Hypocritical Racists

In *Black Innocence and the White Jury*, Professor Sheri Lynn Johnson offers the hypocrisy interpretation of self-evaluation prejudice reports.<sup>22</sup> According to Johnson, "any encouragement that might be drawn from the initial decrease in extreme negative stereotypes must be qualified by the likelihood that newer data reflect some fading of stereotypes—but also some faking."<sup>23</sup> From this viewpoint, prejudice has not decreased nearly as much as it seems; it has just become less socially acceptable. Thus, merely to appear socially desirable, many survey respondents profess racial liberalism. Although Johnson does not give a concrete estimate of how much "faking" the newer data reflect, she does suggest that "it now may be quite common to underreport prejudiced attitudes" by faking racial tolerance.<sup>24</sup>

To support this interpretation of the survey literature, Johnson points to the findings of an experiment in which white subjects were asked to report their responses to blacks under a normal (control) condition and under a "bogus pipeline" condition.<sup>25</sup> In the pipeline condition, a researcher wires his subjects to a machine that the subjects believe will give him an accurate physiological measure of (i.e., a pipeline to) their automatic or "covert" reactions.<sup>26</sup> The researchers then asked these subjects to estimate what the machine was telling the experimenter about their uncontrolled responses to blacks, as the experimenter asked them to rate blacks on various personality

---

22. Johnson, *supra* note 15, at 1648-49; Johnson also offers aversive racism theory as an approach to current racial attitudes, *see infra* notes 58-60 and accompanying text.

23. Johnson, *supra* note 15, at 1648.

24. *Id.* at 1650.

25. *Id.* at 1648.

26. Harold Sigall & Richard Page, *Current Stereotypes: A Little Fading, A Little Faking*, 18 J. PERSONALITY & SOC. PSYCHOL. 247, 248 (1971).



traits, such as ignorance, stupidity, honesty and sensitivity.<sup>27</sup> Researchers assumed that these estimations would correspond to the subjects' "honest" beliefs about blacks.<sup>28</sup> Subjects' estimates of their uncontrolled responses to blacks in the pipeline condition were significantly more negative than the responses to blacks reported by subjects who did not believe that their uncontrolled responses were being monitored.<sup>29</sup>

Johnson characterizes these automatic, uncontrolled physiological responses in this experiment as the subjects' "true feelings"<sup>30</sup> and "pure" attitudes.<sup>31</sup> She interprets these findings as proof of prejudice's persistence notwithstanding survey data to the contrary. This interpretation, however, rests on a failure to distinguish between two distinct sources of negative responses to blacks (and other marginalized social groups)—namely, stereotypes and prejudice. Once this critical distinction is understood, it becomes evident that the bogus pipeline results prove only the persistence of stereotypes, not prejudice, and therefore are perfectly consistent with the proposition that prejudice has decreased significantly over the last 40 years.

Stereotypes consist of well-learned sets of associations among groups and traits established in children's memories at an early age, before they have the cognitive skills to decide rationally upon the personal acceptability of the stereotypes.<sup>32</sup> For example, Phyllis Katz reports a chilling case of a 3-year-old child, who upon seeing a black infant said to her mother, "Look mom, a baby maid."<sup>33</sup> By the time the child turned three, before she had developed the cognitive ability to judge the appropriateness of the stereotypic ascription, the associational link between black women and certain social roles was already forged in her memory.<sup>34</sup>

27. *Id.* at 251.

28. *Id.* at 254.

29. *Id.* at 252. The subjects attributed the negative traits of ostentatiousness, laziness, ignorance, physical dirtiness, stupidity, and unreliability *more* often to blacks when they thought the experimenter could monitor their uncontrolled physiological responses. Furthermore, they attributed the positive traits of intelligence, honesty, and sensitivity to blacks *less* often in the pipeline condition. *Id.*

30. Johnson, *supra* note 15, at 1649.

31. *Id.* at 1649-50. Johnson's interpretation of the meaning of these automatic responses reflects the interpretation given to them by the experimenters who designed the study, Harold Sigall and Richard Page. In their words, "[I]t will be obvious by this point that we have elected to interpret the results of the [bogus pipeline] condition as relatively distortion free, as more honest, and as 'truer' than rating-condition responses. Thus, the [bogus pipeline] may be viewed as a lie detection device which facilitates truthful reporting." Sigall & Page, *supra* note 26, at 254.

32. Patricia G. Devine, *Stereotypes and Prejudice: Their Automatic and Controlled Components*, 56 J. PERSONALITY & SOC. PSYCHOL. 5, 6 (1989).

33. Phyllis A. Katz, *The Acquisition of Racial Attitudes in Children*, in TOWARDS THE ELIMINATION OF RACISM 125, 147 (Phyllis A. Katz ed., 1976).

34. Mary Ellen Goodman found that children at ages 3 to 4 already possess racial awareness, and 25% of 4-year-olds expressed strong racial attitudes. MARY ELLEN GOODMAN, RACE AWARENESS IN YOUNG CHILDREN 47, 245, 252-54 (rev. ed. 1964). More recent research confirms that children typically show evidence of racial awareness by age 3 or 4 and that by the time they reach first grade racial awareness is very well-established. See Katz, *supra* note 33, at 125-26; see also Harold M. Proshansky, *The Development of Intergroup Attitudes*, in 2 REVIEW OF CHILD DEVELOPMENT RESEARCH 311, 314-15 (Lois Wladis Hoffman & Martin L. Hoffman eds., 1966). My personal experience with my

In contrast, prejudice consists of derogatory *personal beliefs*.<sup>35</sup> As Anthony R. Pratkanis points out, beliefs are propositions that people endorse and accept as being true.<sup>36</sup> Thus, prejudiced personal beliefs are the "endorsement or acceptance of the content of a negative cultural stereotype."<sup>37</sup> That a person has a negative stereotype established in her memory does not necessarily mean that she endorses that stereotype. As Devine points out, "although one may have *knowledge of a stereotype*, his or her *personal beliefs* may or may not be congruent with the stereotype."<sup>38</sup> For example, if the 3-year-old child described above grows up and decides that the stereotype of a maid is an inappropriate basis for responding to black women, she may experience a fundamental conflict between the previously established stereotype and the more recently established nonprejudiced personal belief. In such a case, her responses to black women and to blacks generally will turn on whether those responses are based on the well-established stereotype or her more recently adopted nonprejudiced beliefs.

Of course, some people's stereotypes and personal beliefs overlap; that is, some people not only have knowledge of the cultural stereotypes from years of socialization, but they endorse and accept them as well. Research literature classifies these individuals as *high prejudiced*.<sup>39</sup> However, many people have thought about the cultural stereotypes, recognized them as inappropriate bases for responding to others, and deliberately rejected them. Researchers refer to these individuals as *low prejudiced*.<sup>40</sup> Although high- and low-prejudiced persons differ in their personal beliefs about blacks, common socialization experiences have firmly entrenched the cultural stereotype of blacks in the memories of both.<sup>41</sup>

The failure to distinguish between stereotypes and prejudiced personal beliefs leads Johnson and other commentators to take an all-or-nothing approach to prejudice: if a person experiences any stereotype-congruent

---

own son locates the age of racial awareness even earlier. When my son was two years old, having just crossed the threshold of intelligible speech, he announced to his mother and me that his own (in my view, gloriously kinky) hair was "not pretty." Then, pointing to our television and the image there of a model sporting cascading waves of decidedly unkinky hair for a shampoo commercial, he said, "her hair pretty . . . mine not pretty."

35. Of course, a person can be prejudiced in favor of a group as well as against one. See GORDON W. ALLPORT, *THE NATURE OF PREJUDICE* 6-7 (1954). Since the focus of this article is the reactions of legal decisionmakers to outgroups, my focus is on prejudice against a group.

36. Anthony R. Pratkanis, *The Cognitive Representation of Attitudes*, in *ATTITUDE STRUCTURE AND FUNCTION* 71, 91 (Anthony R. Pratkanis et al. eds., 1989).

37. Margo J. Monteith et al., *Prejudice and Prejudice Reduction: Classic Challenges, Contemporary Approaches*, in *SOCIAL COGNITION: IMPACT ON SOCIAL PSYCHOLOGY* 323, 333-34 (Patricia G. Devine et al. eds., 1994).

38. Devine, *supra* note 32, at 5.

39. See, e.g., Patricia G. Devine et al., *Prejudice With and Without Compunction*, 60 *J. PERSONALITY & SOC. PSYCHOL.* 817, 817-19 (1991).

40. See, e.g., *id.*

41. For a discussion of empirical evidence that the black stereotype is equally well-established in the memories of both high- and low-prejudiced persons, see *infra* text accompanying notes 118-122.

responses in any situation, she is prejudiced. This view fails to recognize that a change in a person's beliefs does not instantly extinguish habitual responses derived from well-learned stereotypes. Because stereotypes are established in children's memories at an early age and constantly reinforced through the mass media and other socializing agents, stereotype-congruent responses may persist long after a person has sincerely renounced prejudice. As Patricia Devine et al. point out, nonprejudiced beliefs and stereotype-congruent thoughts and feelings may coexist within the same individual.<sup>42</sup> Dr. Thomas Pettigrew, a leading authority on stereotypes and prejudice, has described one example of this conflict: "Many Southerners have confessed to me . . . that even though in their minds they no longer feel prejudice against blacks, they still feel squeamish when they shake hands with a black. The feelings are left over from what they learned in their families as children."<sup>43</sup>

There is strong empirical evidence that the vast majority of low-prejudiced people realize that they are prone to stereotype-congruent responses, i.e., that their *actual* reactions to out-group members sometime conflict with their personal standards for how they *should* respond. In one recent study, researchers gave a sample of several hundred white subjects (college students, very few of whom were high-prejudiced)<sup>44</sup> a questionnaire, the first section of which asked them to report their personal standards for how they *should* respond in five different situations involving black people.<sup>45</sup> For example, one situation read as follows: "Imagine that a Black person boarded a bus and sat next to you. You *should* feel uncomfortable that a Black is sitting next to you."<sup>46</sup>

The subjects were asked to circle the number between 1 (strongly disagree) and 7 (strongly agree) that best reflected their personal standard for

---

42. Devine et al., *supra* note 39, at 817.

43. Quoted in Daniel Coleman, "Useful" Modes of Thinking Contribute to the Power of Prejudice, N.Y. TIMES, May 12, 1987, at C1, C10. Gordon Allport, another leading authority on stereotypes and prejudice, also has noted the conflict between ingrained stereotypes and new non-prejudiced beliefs. See ALLPORT, *supra* note 35, at 328 ("Defeated intellectually, prejudice lingers emotionally.").

44. The subjects' prejudice level was determined on the basis of their responses to a seven-item Modern Racism Scale, a nonreactive measure of negative attitudes towards blacks. Devine et al., *supra* note 39, at 819. "The Modern Racism Scale has proven to be useful in predicting a variety of behaviors including voting patterns and reactions to busing." Devine, *supra* note 32, at 7.

45. Devine et al., *supra* note 39 at 819.

46. *Id.*

Two of the remaining four situations also focused on feelings subjects could have in response to situations involving Black people. One situation involved feeling upset that a Black couple moved in next door. The other involved feeling uncomfortable that a job interviewer is Black. The final two situations focused on stereotypic thoughts subjects might have in contact situations with Black people. One thought situation involved seeing three middle-aged Black men on a street corner in the afternoon and thinking "Why don't they get a job?" The other thought situation involved seeing a black woman with several small children and thinking "How typical."

*Id.*

how they *should* respond in each situation. The second section of the questionnaire asked the subjects to report on the 1-to-7 scale how they believed they actually *would* respond in the same five situations. Out of the 101 cases, 71% of the subjects reported actual *would* responses that were more negative than their *should* responses, which reflected their personal standards for how they should respond.<sup>47</sup> Separate studies found similar should-would discrepancies in responses to homosexual men.<sup>48</sup> These studies also investigated whether the subjects' personal standards (shoulds) were well-internalized (i.e., viewed by the subjects as highly important and as central to their personal identity or merely derived from society's standards).<sup>49</sup> Researchers found that low-prejudiced subjects strongly internalized their personal standards,<sup>50</sup> and that these subjects felt compunction (guilt and self-criticism) when they transgressed the standards.<sup>51</sup>

These findings, which several later studies have replicated,<sup>52</sup> suggest a less pessimistic interpretation of the bogus pipeline results than Johnson adopts. That a subject reports more negative responses about blacks when he believes an experimenter can monitor his autonomic nervous system (or what Johnson refers to as his "true attitudes") does not prove that he is truly prejudiced or that he is faking his more positive responses on questionnaires. It may show only that he realizes, as most low-prejudiced people do, that he is prone to stereotype-congruent responses. Although he may not endorse these responses and may feel compunction about experiencing them, he may believe that the pipeline will detect their presence. In other situations, however, such as responding to a questionnaire, the low-prejudiced person may inhibit his stereotype-congruent responses and replace them with responses based on his nonprejudiced personal standards. A model of prejudice that recognizes the distinction between stereotypes and prejudiced personal beliefs—a model I shall call the "dissociation model"—points to the possibility of inhibiting and replacing stereotype-congruent responses with nonprejudiced responses derived from nonprejudiced personal beliefs. If nonprejudiced personal beliefs can counteract stereotypes in this way, perhaps there is hope for combatting the influence of ubiquitous derogatory stereotypes.

Studies support the proposition that responses derived from nonprejudiced personal beliefs can inhibit and replace responses derived from ste-

---

47. *Id.* at 820.

48. *Id.* at 822.

49. *Id.* at 824-27.

50. *Id.* at 826.

51. *Id.* at 827.

52. See, e.g., Margo J. Monteith et al., *Self-Directed Versus Other-Directed Affect as a Consequence of Prejudice-Related Discrepancies*, 64 J. PERSONALITY & SOC. PSYCHOL. 198, 200-08 (1993) (hereinafter Monteith, *Self-Directed*); Margo J. Monteith, *Self-Regulation of Prejudiced Responses: Implications for Progress in Prejudice-Reduction Efforts*, 65 J. PERSONALITY & SOC. PSYCHOL. 469, 471-78 (1993) (hereinafter Monteith, *Self-Regulation*).

reotypes. Research has demonstrated that low- and high-prejudiced people are equally prone to stereotype-congruent responses when they cannot consciously monitor their responses to questions.<sup>53</sup> However, low- and high-prejudiced people have given very different responses when they have had to think consciously about what their responses imply about their self-image. For example, one study asked subjects to list all of their own thoughts (e.g., beliefs, feelings, expectancies) about blacks under strictly anonymous conditions, thus eliminating any reason to manufacture "correct" responses.<sup>54</sup> Researchers found that the high-prejudiced subjects listed primarily negative stereotypical thoughts about blacks and were inclined to stereotype.<sup>55</sup> In contrast, the low-prejudiced subjects wrote few pejorative thoughts; they reported beliefs that contradicted the stereotype and emphasized the importance of racial equality.<sup>56</sup> These results make intuitive sense. For low-prejudiced people, writing stereotype-congruent thoughts would contradict their personal beliefs and threaten their nonprejudiced identity. But because beliefs of high-prejudiced people overlap with stereotypes, conscious reflection should not inhibit their stereotype-congruent responses. Thus, if personal beliefs *really matter*, if they can counteract the stereotype-congruent responses to which research shows high- and low-prejudiced people are equally prone, then the thoughts that low-prejudiced subjects anonymously list about blacks should be very different from the thoughts anonymously listed by high-prejudiced subjects.

The findings of this thought-listing study, which are strongly confirmed and extended by other research that I discuss below,<sup>57</sup> reveal that much more than semantics is at stake in the distinction between stereotypes and prejudice. For inasmuch as negative stereotypes and personal beliefs diverge, as they do in low-prejudiced people, they imply different responses to stereotyped groups. This insight enables us to investigate the interplay between the two conceptually distinct sets of responses, and to develop strategies for activating the responses based on nonprejudiced personal beliefs and inhibiting the stereotype-congruent responses. However, before elaborating a framework for working out the full implications of the interplay between stereotypes and nonprejudiced personal beliefs, I will consider the other major attack on the validity and efficacy of nonprejudiced personal beliefs: aversive racism.

---

53. See *infra* text accompanying notes 118-122.

54. Devine, *supra* note 32, at 12-13.

55. *Id.* at 14-15.

56. *Id.* at 14.

57. See *infra* text accompanying notes 124-130.

### B. Aversive Racists

The dominant model of prejudice in current legal literature is the theory of aversive racism.<sup>58</sup> Whereas the hypocritical racist model posits that people who express nonprejudiced personal beliefs are manipulating their self-presentation to appear more socially desirable, the aversive racist model holds that ostensibly nonprejudiced people are not so much deceiving others as fooling themselves.<sup>59</sup> The two models are not mutually exclusive, but complementary. Commentators freely switch from one to the other<sup>60</sup> by dismissing the validity of people's racially liberal self-descriptions and nonprejudiced personal beliefs.

The theory of aversive racism begins with the proposition that most Americans are highly committed to egalitarian values. Therefore, they desire to maintain an egalitarian, nonprejudiced self-image. This desire causes them to express nonprejudiced personal beliefs.<sup>61</sup> Such professed nonprejudiced beliefs are not to be confused with genuine—i.e., well-internalized—nonprejudiced beliefs, for deep down “aversive racist[s] believe[] in white superiority”<sup>62</sup> and “do not want to associate with blacks.”<sup>63</sup> Desperately clinging to their egalitarian, nonprejudiced values and self-image, aversive racists repress their negative feelings and beliefs about blacks. Lawrence refers to these repressed antiblack beliefs as “hidden prejudice,”<sup>64</sup> and offers a Freudian theory of unconscious motivation to explain how “we all harbor prejudiced attitudes that are kept from our consciousness.”<sup>65</sup> Moreover, since aversive racists do not recognize their antiblack attitudes, the prospects for prejudice reduction are particularly dim. Here the pessimism of the aversive racism model asserts itself. Writing from this perspective, Professor Peggy Davis observes, “[i]t is diffi-

58. E.g., Davis, *supra* note 13, at 1564-65; Johnson, *supra* note 15, at 1649; Lawrence, *supra* note 13, at 335-36.

59. Johnson, *supra* note 13, at 1029-30 (describing denial as a central property of unconscious racism).

60. In *Black Innocence and the White Jury*, for example, Johnson argues that the perceived faking of nonprejudiced responses revealed in the “bogus pipeline” experiment is “complemented by observations concerning the prevalence of two kinds of racism. Dominative racists express their bigoted beliefs openly . . . while aversive racists do not want to associate with blacks but do not often express this feeling.” Johnson, *supra* note 15, at 1649. She notes that social scientists now believe that “aversive manifestations of racism increasingly predominate in all parts of the country.” *Id.* Taken together, the dominative racist, the pseudoliberal, and the aversive racist make up the taxonomy of racists that frames the discussion of prejudice among legal scholars. And, inasmuch as “we are all racists,” presumably we all fit into one of these subcategories. Alternatively, since proponents of the proposition that “we are all racists” at least recognize their racism, perhaps another category must be added to the current taxonomy of racists, such as the “enlightened racist.”

61. Davis, *supra* note 13, at 1564-65; Samuel L. Gaertner & John F. Dovidio, *The Aversive Form of Racism*, in PREJUDICE, DISCRIMINATION, AND RACISM 61, 84 (John F. Dovidio & Samuel L. Gaertner eds., 1986).

62. Lawrence, *supra* note 13, at 335.

63. Johnson, *supra* note 15, at 1649.

64. Lawrence, *supra* note 13, at 335.

65. *Id.* at 339.

cult to change an attitude that is unacknowledged. Thus, 'like a virus that mutates into new forms, old-fashioned prejudice seems to have evolved into a new type that is, at least temporarily, resistant to traditional . . . remedies.'<sup>66</sup>

Commentators contend that proof of aversive racism lies in the discrepancy between responses to blacks that are consciously monitored and those that are not consciously monitored. Whenever aversive racists consciously monitor their responses to blacks, they do not discriminate against them since discrimination would undermine their egalitarian self-images. For example, verbal responses to questionnaires designed to measure racial prejudice can be monitored consciously by the respondents and therefore cannot identify aversive racists.<sup>67</sup> More generally, if the situation clearly calls for a nonprejudiced response, or if a *nonracial* justification or rationalization for engaging in a prejudiced response cannot be generated, the response will be positive because it cannot escape being consciously monitored.

In contrast, when the situation is normatively ambiguous, or when a *nonrace-related* justification is handy, the covert antiblack attitudes and beliefs of aversive racists find expression in racial discrimination.<sup>68</sup> For example, white research subjects led to believe that a person was in distress helped black victims as often as white victims when there was no ostensible justification for a failure to help.<sup>69</sup> However, if the subjects knew that someone was available to help, they "helped black victims much less frequently than they helped white victims (38% vs. 75%)."<sup>70</sup> According to proponents of the aversive racism model, the availability of other potential rescuers provided subjects with a convenient nonracial excuse for not helping the black victims. This interpretation of the helping behavior study carries very discouraging implications for racially fair dispute resolution. For in formal legal proceedings, finding a nonracial reason to discriminate against a black litigant is especially easy to do—one simply gives more weight to the evidence favoring the opposing litigant.

The aversive racism model, however, is empirically and conceptually incomplete. One empirical problem with the model concerns its assumption that aversive racists—who, according to commentators, now include most

66. Davis, *supra* note 13, at 1565 (quoting Gaertner & Dovidio, *supra* note 61, at 85-86).

67. Gaertner & Dovidio, *supra* note 61, at 67 ("Given the high salience of race and racially symbolic issues on questionnaires designed to measure racial prejudice, as well as aversive racists' vigilance and sensitivity to these issues, effective questionnaire measures of aversive racism, in our opinion, would be difficult if not impossible to develop.")

68. See Johnson, *supra* note 13, at 1030-31.

69. Gaertner & Dovidio, *supra* note 61, at 77. In fact, white rescuers helped blacks *more* often than whites (94% to 81%). *Id.*

70. *Id.* In addition, when the subjects believed a bystander could intervene, they also showed "lower levels of arousal with black than with white victims," as measured by their heartrate escalation when the accident occurred to the victim. *Id.* at 78.

Americans<sup>71</sup>—are not aware of their conflicting reactions to blacks; their antiblack thoughts and feelings are supposedly excluded from consciousness. If this assumption were accurate, most survey respondents would not report discrepancies between their standards for how they should respond to blacks and how they actually would respond, since they are unaware. Yet the vast majority of subjects in several studies recognized and acknowledged that they sometimes experience such discrepancies.<sup>72</sup> Thus, although the aversive racism framework may describe some white Americans, it almost certainly does not account for most.

Another empirical problem with the aversive racism theory concerns the Freudian theory of unconscious motivation to which it is often wedded.<sup>73</sup> As Lawrence frankly admits, psychoanalytic theory presents real difficulties for empirical verification.<sup>74</sup> A model, such as aversive racism, whose theoretical underpinnings are not empirically demonstrable demands an intellectual leap of faith that many may be unwilling to make. Alternatively, commentators have attempted to explain aversive racism by drawing on research in cognitive psychology.<sup>75</sup> Although this research, unlike Freudian psychoanalysis, provides a rich lode of empirical findings on the inherent tendency of the human mind to prejudge and overgeneralize (a lode I mine actively in Part II of this article), commentators writing from this perspective never adequately explain how the prejudgments and overgeneralizations resulting from this tendency escape a person's awareness. On the contrary, commentators suggest that we are aware of the resulting stereotypes, but that we experience them as rational reflections of objective reality rather than as figments of our distorted cognitive processes.<sup>76</sup> Thus, neither the Freudian nor the cognitive framework gives

---

71. This is the implication of Professor Lawrence's statement that "we are all racists . . . [although] most of us are unaware of our racism," Lawrence, *supra* note 13, at 322, in combination with his suggestion that "this hidden prejudice has become the more prevalent form of racism," *id.* at 335. Professor Johnson also states that "aversive manifestations of racism increasingly predominate in all parts of the country." Johnson, *supra* note 15, at 1649.

72. See *supra* text accompanying notes 44-51.

73. E.g., Davis, *supra* note 13, at 1562 (describing the projection of forbidden prejudices as a psychological defense mechanism); Lawrence, *supra* note 13, at 331-36 (explaining racism's relationship to Freud's psychoanalytic concepts of the Ego and the Id).

74. Lawrence, *supra* note 13, at 331 n.55.

75. *Id.* at 336-39; Davis, *supra* note 13, at 1561-62.

76. Writing from a cognitive psychology perspective, Peggy C. Davis argues that a person who she describes as reacting negatively to blacks has "assimilated negative stereotypes about blacks before she reached the age of judgment. She will, therefore, have accepted them as truth rather than opinion." Davis, *supra* note 13, at 1562. Moreover, although Charles R. Lawrence III elegantly explains how people lack awareness of the ways in which they selectively *process* social information to form and reinforce stereotypes, he does not explain how people can also be unaware of the *presence* of these stereotypes in their thought processes once they have been formed:

While the individual may be aware of the selectively perceived facts that support his categorization or simplified understanding, he will not be aware of the process that has caused him to deselect the facts that do not conform with his rationalization. Thus, racially prejudiced behavior that is actually the product of learned cultural preferences is experienced as a reflection of rational deduction from objective observation, which is nonprejudicial



an empirically grounded account of how factfinders could fall into negative responses to blacks and members of other stereotyped groups without being aware of those responses as they are occurring. To convince courts to factor psychological propositions about unconscious discrimination into their formulation and application of legal rules, it is important to frame a model of such discrimination that is firmly rooted in empirical research.

The conceptual problem with the prevailing aversive racism model concerns its tendency to conflate stereotype and prejudice. Recall Lawrence's often quoted assertion that "[t]o the extent that this cultural belief system has influenced all of us, we are all racists."<sup>77</sup> Since ethnic attitudes and stereotypes are part of a society's social heritage and no one can escape learning the prevailing attitudes and stereotypes assigned to the major ethnic groups,<sup>78</sup> Lawrence's position states that we are all racists. Lawrence does not, however, explore the implications of the fact that people do not always endorse the knowledge structures that socialization has established in their memories. For example, although socializing forces undoubtedly have entrenched the cultural stereotype of women in the memory of feminists as well as every other American, feminists could be called "sexists" only in a Pickwickian sense. One reason it seems so anomalous to apply the value-laden term "sexist" to feminists is because feminists have both renounced the cultural stereotype about women and developed egalitarian personal beliefs about women. Thus, feminists have two distinct and conflicting cognitive structures concerning women: the cultural stereotype and their egalitarian personal beliefs. Similarly, low-prejudiced people have two conflicting cognitive structures concerning blacks: the black cultural stereotype and their nonprejudiced personal beliefs. Calling feminists "sexists" and low-prejudiced persons "racists" identifies them more with the well-learned cultural stereotype than with their personal beliefs, and implies that the stereotype is somehow the more compelling of the two knowledge structures.

Instead of debating which cognitive structure—stereotypes or nonprejudiced personal beliefs—is *the* defining feature of a low-prejudiced person's mental processes, or which is the more important determinant of her responses to members of stereotyped groups, I argue that both types of responses occur and that both can influence decisionmakers' behavior. In developing this analysis, I draw extensively on recent research in social cognition, a defining characteristic of which is its emphasis on the type of social information that is stored in memory. Following Devine et al., I

---

behavior. The decisionmaker who is unaware of the selective perception that has produced her stereotype will not view it as a stereotype. She will believe that her actions are motivated not by racial prejudice but by her attraction or aversion to the attributes she has "observed" in the groups she has favored or disfavored.

Lawrence, *supra* note 13, at 339.

77. See *supra* text accompanying note 19.

78. EHRLICH, *supra* note 18, at 110.

argue that stereotypes and personal beliefs (or attitudes) represent conceptually distinct and potentially conflicting subsets of information about ethnic or racial groups, and that a different cognitive process governs each distinct subset of information. Further, understanding the interplay between these different types of information and processes both accounts for unconscious discrimination and suggests strategies for combatting it. Perhaps most importantly, I will try to ground each step of my analysis of unconscious discrimination and discrimination reduction techniques in empirically demonstrated propositions. Thus, I hope to lay down a solid empirical foundation for the claim that unconscious discrimination routinely infects legal decisionmaking, without ignoring the utility of nonprejudiced personal standards in the fight against such discrimination.

## II

### AN EMPIRICALLY DEMONSTRABLE MODEL OF UNCONSCIOUS DISCRIMINATION

[G]iven a sensory input with equally good fit to two nonoverlapping categories, the more accessible of the two categories would "capture" the input.<sup>79</sup>

It is widely believed that our judgments and memories of others turn on whatever information about them has been made available to us. But if information alone were sufficient to determine our social judgments, then reasonable people who are exposed to the same information about someone should form the same judgments. Yet, people often form different judgments and recollect different facts, even when exposed to the same information. Thus, in addition to information from the environmental and social context, the perceiver's cognitive structures and processes must also determine his or her social judgments. The following question therefore arises: What are these processes and what implications do they carry for social judgments of blacks and other stereotyped groups?

Following Jerome Bruner,<sup>80</sup> social cognition researchers conceptualize the process that underlies the perception of persons as a categorization process. In Lawrence's apt summary of this perspective, "[a]ll humans tend to categorize in order to make sense of experience. Too many events occur daily for us to deal successfully with each one on an individual basis; we must categorize in order to cope."<sup>81</sup> Thus, a person who is asked to judge another's behavior must first take whatever information she receives about the other's behavior and interpret, or *encode*, this behavior by assigning it to a category. According to social psychologists E. Tory Higgins and Gillian King, social and personal categories include information about

---

79. Jerome S. Bruner, *On Perceptual Readiness*, 64 *PSYCHOL. REV.* 123, 132 (1957).

80. *See id.*; Jerome S. Bruner, *Social Psychology and Perception*, in *READINGS IN SOCIAL PSYCHOLOGY* (Eleanor E. Maccoby et al. eds., 3d ed. 1958).

81. Lawrence, *supra* note 13, at 337.

social groups (e.g., blacks, women, gays and lesbians), social roles and occupations (e.g., spouses, maids, police officers), traits and behaviors (e.g., hostile, crime-prone, patriotic, and intelligent), and social types (e.g., intellectual, social activists, and rednecks).<sup>82</sup> Once the behavior is assigned to one of these categories, it is stored in memory, from which it subsequently can be retrieved to make further inferences and predictions about the person.

When individuals must judge another's behavior, however, they are unlikely to perform an exhaustive search of memory for all potentially relevant categories, compare the behavior to each such category, and then characterize the behavior in terms of the category with the best "fit." Rather, they are likely to base their judgment on the category that happens to be the most readily *accessible* at the time the information is received.<sup>83</sup> Steven L. Neuberg explains this phenomenon with the following example:

[J]ust after viewing an extremely violent film in which a heartless mugger preys upon innocent travelers of the city streets, a moviegoer would have a greater than usual tendency to perceive the behavior of a stranger who bumps into him or her as reflecting hostility or aggressiveness. Alternatively, after viewing a comedy featuring the inept Inspector Clouseau, the moviegoer might be more likely to perceive the identical social interaction in terms of the stranger's clumsiness. In each example, the film preceding the interaction "primed" particular cognitive categories that subsequently influenced the interpretation of the incident.<sup>84</sup>

Numerous studies confirm this intuitive account of the centrality of category accessibility in social perceptions and judgments. In one classic study, E. Tory Higgins, William E. Rholes, and Carl R. Jones posited that unobtrusively exposing subjects to certain personality trait terms in one exercise would activate, or *prime*, the categories to which these terms referred, making it more likely the subjects would use the categories to characterize a person in an unrelated context.<sup>85</sup> To test this hypothesis, Higgins, Rholes, and Jones asked subjects to perform a complex cognitive task that momentarily exposed them to several trait terms.<sup>86</sup> Later, in what

---

82. E. Tory Higgins & Gillian King, *Accessibility of Social Constructs: Information-Processing Consequences of Individual and Contextual Variability*, in PERSONALITY, COGNITION, AND SOCIAL INTERACTION 69, 71-72 (Nancy Cantor & John F. Kihlstrom eds., 1981).

83. Steven L. Neuberg, *Behavioral Implications of Information Presented Outside of Conscious Awareness: The Effect of Subliminal Presentation of Trait Information on Behavior in the Prisoner's Dilemma Game*, in 6 SOCIAL COGNITION 207, 208 (1988).

84. *Id.*

85. E. Tory Higgins et al., *Category Accessibility and Impression Formation*, 13 J. EXPERIMENTAL SOC. PSYCHOL. 141, 141-45 (1977).

86. *Id.* at 145-46. Some subjects were momentarily exposed to favorable trait terms (e.g., "adventurous"), some to unfavorable trait terms (e.g., "reckless"), and some to trait terms that were inapplicable to the behavior of the protagonist in the passage the subjects were about to read (e.g., "obedient" or "disrespectful"). *Id.* at 145.

ostensibly was an unrelated experiment on reading comprehension, the subjects read a paragraph about a target person, which was ambiguous as to his likability.<sup>87</sup> After reading the passage, subjects characterized the target person in their own words. As predicted, subjects unobtrusively exposed to favorable trait terms tended to use those terms or their synonyms in characterizing the target, while subjects exposed to unfavorable terms tended to use those terms or their synonyms in their characterizations.<sup>88</sup> In contrast, control subjects that researchers exposed to trait terms that were not applicable to interpreting the target's behavior did not vary systematically in their characterizations.<sup>89</sup>

These results carry enormous implications for judgments and evaluations of stereotyped groups. If cues of group membership such as race serve to prime trait categories such as hostility, people will systematically view behaviors by members of certain racial groups (e.g., blacks) as more menacing than the *same* behaviors by members of other racial groups (e.g., whites).

Birt L. Duncan's research provides background for understanding these implications. Duncan found that whites interpreted the same ambiguous shove as hostile or violent when the actor was black and as "playing around" or "dramatizing" when the actor was white.<sup>90</sup> He assumed that

---

87. *Id.* at 145. Specifically, the paragraph was ambiguous with respect to several personality traits. For example, the protagonist of the paragraph was described as thinking about crossing the Atlantic in a sailboat, behavior that could be characterized favorably as "adventurous" or unfavorably as "reckless." *Id.*

88. *Id.* at 147-50.

89. *Id.*

90. 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, 595-97 (1976). Duncan had 96 white undergraduates individually rate a series of interactions between two "other subjects" that culminated in an ambiguous shoving event. The two "other subjects" (both male) were actually confederates acting out a script. The experimental session consisted of a videotape of the two actors discussing a hypothetical problem; however, the subject who observed the tape was led to believe that the discussion actually was taking place in another room. *Id.* at 592. The subject was asked to evaluate the behavior of the "actors" six times at precise intervals, which the experimenter signaled to him during the tape. *Id.* To evaluate the actors' behavior, the subjects had to fit the behavior into one of ten major categories on a rating form. The ten major categories were *dramatizes, gives information, gives opinion, gives suggestion, asks for information, asks for opinion, asks for suggestion, playing around, aggressive behavior, and violent behavior*. *Id.* at 594. The subjects' final evaluations—their sixth ratings—were designed to coincide with the heated discussion and ambiguous shove near the end of the interactions; thus, this sixth rating was the major dependent measure. *Id.* at 592.

The major independent variables were the racial identities of the actor who initiated the ambiguous shove (the "protagonist") and the actor who received the shove (the "victim"). *Id.* The subjects were randomly assigned to one of the following experimental conditions (tapes): black protagonist-white victim; white protagonist-black victim; black protagonist-black victim; white protagonist-white victim. *Id.* at 592-94. The results of this experiment are disturbing and unequivocal. When the protagonist was black and the victim white, 75% of the subjects characterized the ambiguous shove as "violent behavior," whereas when the protagonist was white and the victim black, only 17% so characterized it. On the other hand, 42% of the subjects perceived the shove as "playing around" or "dramatizing" when the protagonist was white and victim black, compared with only 6% in the black-protagonist/white-victim conditions. *Id.* at 595. The discrepancy between white-protagonist/white-victim condition and

category accessibility best explains this differential perception of violence as a function of the protagonist's race.<sup>91</sup> Duncan assumed that the presence of the black actor primed the stereotype of blacks, and since the stereotype associates blacks with violence, the violent behavior category was more accessible when interpreting behavioral information about blacks than whites.<sup>92</sup> Sager and Schofield replicated Duncan's findings in studies of schoolchildren.<sup>93</sup> They found that both black and white children rated ambiguously aggressive behaviors (e.g., bumping in the hallway) of black actors as being more mean or threatening than the same behaviors of white actors.<sup>94</sup>

Although Duncan's study provides compelling evidence that race influences category accessibility, it does not determine whether the influence is unconscious or conscious. It is possible that upon noticing the racial identity of the black protagonist, the subjects (or some percentage of them) formed a conscious expectation for instances of trait categories stereotypically associated with blacks (e.g., hostile, prone to violence). Indeed, research indicates that expecting to see an instance of a trait category increases the likelihood that a person will process ambiguous information by putting it into that category.<sup>95</sup>

On the other hand, Duncan's subjects (or some percentage of them) could have been sincerely nonprejudiced and refrained from consciously forming any race-based expectation of hostility, yet the mere presence of the black protagonist may have automatically (i.e., unconsciously) activated the black stereotype, including the hostility trait category that figures so prominently in that stereotype.<sup>96</sup> Thus, the subjects could have sincerely

---

black-protagonist/black-victim condition is also drastic: 69% of the subjects perceived the within-group (black-black) condition as violent compared with 13% in the white-white conditions. *Id.* Thus, the subjects in this experiment were much more likely to characterize an act as violent when it was performed by a black than when the same act was committed by a white. *Id.*

91. *Id.* at 591.

92. *Id.*

93. H. Andrew Sager & Janet Ward Schofield, *Racial and Behavioral Cues in Black and White Children's Perceptions of Ambiguously Aggressive Acts*, 39 J. PERSONALITY & SOC. PSYCHOL. 590, 594-97 (1980).

94. *Id.*

95. "The accessibility of a construct will increase when the estimate of the likelihood of occurrence of a construct instance increases." Higgins & King, *supra* note 82, at 75. For example, in Harold Kelley's study of labeling effects on impression formation, students' ratings of a new instructor were more favorable when the experimenter described the instructor as a "warm" person to the students before the instructor's arrival than when the experimenter described the instructor as a "cold" person. Harold H. Kelley, *The Warm-Cold Variable in First Impressions of Persons*, 18 J. PERSONALITY & SOC. PSYCHOL. 431, 435 (1950). According to Kelley, the experimenter's prior description prepared the students to expect instances of the category designated by the label as well as instances of other categories assumed to be closely related to this category (e.g., "friendly" and "helpful"). *Id.* at 435-36.

96. This is also another possible interpretation of the results of Kelley's study—that the mere exposure to the label "warm" may have activated the trait category designated by the label. Higgins & King, *supra* note 82, at 73-74. This passive priming effect would have constituted what I shall refer to later in the article as an "automatic process." See *infra* text accompanying notes 102-107. Furthermore, because both processes—the controlled and the automatic—could have simultaneously operated on the

renounced racial prejudice and still unconsciously practiced discrimination against the black actor. To understand how a knowledge structure such as a stereotype can operate outside a person's awareness and determine his or her responses to others, it is necessary to understand the distinction between habits and decisions, a distinction cognitive psychologists characterize in terms of *automatic* versus *controlled* processes.<sup>97</sup> Since this distinction also sheds light on the interplay of stereotypes and personal beliefs in responses to members of stereotyped groups, and ultimately points to strategies for discrimination reduction, I will discuss it in detail.

A habit is "an action that has been done many times and has become automatic. That is, it is done without conscious thought."<sup>98</sup> In contrast, a decision to take or not to take an action involves conscious thought.<sup>99</sup> The distinction between habit and conscious decision is one of the oldest concepts in psychology. In his *Principles of Psychology*, William James described the origins and consequences of habit as follows:

[A]ny sequence of mental action which has been frequently repeated tends to perpetuate itself; so that we find ourselves automatically prompted to *think, feel, or do* what we have been before accustomed to think, feel, or do, under like circumstances, without any consciously formed *purpose*, or anticipation of results.<sup>100</sup>

James concluded that it is necessary to free limited consciousness from the many mundane requirements of life by removing frequently used or habitual mental sequences from conscious awareness.<sup>101</sup>

The current model of habits and decisions employed by cognitive psychologists is not appreciably different from that outlined by James over a century ago, except that the current model expresses the distinction in terms of *automatic* versus *controlled* processes.<sup>102</sup> According to D.L. Ronis et al., "[h]abits are the results of automatic cognitive processes."<sup>103</sup> As Patricia G. Devine points out, "[a]utomatic processes involve the unintentional or spontaneous activation of some well-learned set of associations or

---

same underlying category ("warm" or "warm disposition"), the two processes were probably mutually facilitative, such that both served to make the underlying category more accessible. What happens, however, when the two processes are mutually antagonistic, i.e., when the controlled processes activate one category or knowledge structure while the automatic processes activate a very different category or knowledge structure? Which process determines which knowledge structure the perceiver will use to judge another person? I address these critical questions in Part III.

97. See *infra* text accompanying notes 102-105.

98. David L. Ronis et al., *Attitudes, Decisions, and Habits as Determinants of Repeated Behavior*, in *ATTITUDE STRUCTURE AND FUNCTION* 213, 218 (Anthony R. Pratkanis et al. eds., 1989).

99. *Id.*

100. 1 WILLIAM JAMES, *PRINCIPLES OF PSYCHOLOGY* 112 (1890) (quoting WILLIAM B. CARPENTER, *PRINCIPLES OF MENTAL PHYSIOLOGY* 339-45 (1874)).

101. *Id.* at 112-14. Thus, like the need to think in terms of simplifying categories, the mind's need to form habits arises from the fact that it is a limited capacity processor.

102. Walter Schneider & Richard M. Shiffrin, *Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention*, 84 *PSYCHOL. REV.* 1, 2-3 (1977).

103. Ronis et al., *supra* note 98, at 219.

responses that have been developed through repeated activation in memory."<sup>104</sup> Controlled processes, on the other hand, "are intentional and require the active attention of the individual."<sup>105</sup> Learning to drive a car provides a useful illustration of this distinction. When you first get behind the wheel, virtually every maneuver is a controlled response. Deciding when and how to apply your foot to the pedals as you turn the steering wheel or manually shift gears demands concentration and effort. After enough practice, however, these maneuvers become automatic. You can accelerate, brake, and steer while contemplating health care reform or talking to a traveling companion. The well-learned motor responses occur without conscious effort.

A critical characteristic of habits or automatic processes is that they can operate independently of conscious decisions to break with old patterns of responses and adopt new ones.<sup>106</sup> Thus, attitudes and beliefs can change without a corresponding change in established habits, resulting in a conflict between currently endorsed responses and old habitual responses.<sup>107</sup> Anyone who has ever tried to break a bad habit knows the persistence of habitual responses in the face of decisions to adopt new ones.

Applied to the relationship between stereotypes and personal beliefs, the habit-decision/automatic-controlled processes distinction provides critical theoretical support for understanding the *more* and *less* conscious aspects of responses to blacks (and members of other stereotyped groups). As discussed earlier, the black stereotype is established in children's memories before children develop the cognitive ability to critically evaluate and decide on the stereotype's acceptability.<sup>108</sup> Further, the social environ-

---

104. Devine, *supra* note 32, at 6.

105. *Id.*

106. Ronis et al., *supra* note 98, at 220-22.

107. *Id.*

108. See *supra* text accompanying notes 32-34.

ment,<sup>109</sup> including the mass media,<sup>110</sup> incessantly reactivates this stereotype. Thus, the stereotype is an ingrained set of associations (i.e., a habit) that involves automatic processes. Nonprejudiced personal beliefs, on the other hand, are necessarily newer cognitive structures that result from a low-prejudiced person's conscious *decision* that stereotype-based responses to blacks are unacceptable.<sup>111</sup>

It follows that these decisions to renounce the already-established stereotype do not come to mind (i.e., are not reactivated) nearly as frequently as the social environment automatically activates the stereotype. Because the stereotype has a longer history and greater frequency of activation than the more recently acquired personal beliefs, even people with well-internalized nonprejudiced beliefs are likely to experience a fundamental conflict between the stereotype and their personal beliefs. The discrepancies that most low-prejudiced subjects report between how they believe they *should* respond and how they actually *would* respond in contact situations with blacks (as well as gays) reflect this conflict.<sup>112</sup> That these subjects also report feeling compunction (i.e., guilt and self-criticism) as a result of these discrepancies<sup>113</sup> implies that they regard the stereotype-congruent responses as essentially a bad habit.

This analysis assumes that just as habitual responses (like putting on a seat belt) may be triggered automatically by the presence of relevant environmental cues (like sitting in a car),<sup>114</sup> stereotype-congruent responses may be triggered automatically by a group membership cue such as a per-

---

109. The content of a stereotype may be based partly on the unequal distribution of members of different groups in different social roles. Researchers have found that observing members of different groups in different roles can influence the actual content of stereotypes. See Alice H. Eagly & Valerie J. Steffen, *Gender Stereotypes Stem From the Distribution of Women and Men Into Social Roles*, 46 J. PERSONALITY & SOC. PSYCHOL. 735, 751-52 (1984). A prominent social psychologist who spent a year as a high-school exchange student in the Republic of South Africa described how an observer in that country could learn negative black stereotypes from the social environment:

Groups of supervised black prisoners were often seen in local parks and on municipal property performing gardening and maintenance tasks. White prisoners, on the other hand, were not required to perform these tasks in public. What might be the implication of seeing only black prisoners fill these roles? First, one might conclude that blacks are more likely to engage in criminal behavior than whites. Whereas relatively few white prisoners were ever seen in public, black convicts were regularly on display in the community. Second, one might conclude that blacks are particularly adept at tasks such as gardening and maintenance, given that they are seen performing these tasks every day. Thus, the white South Africans' stereotype that blacks are criminally prone and capable primarily of certain types of labor may arise from seeing them on a daily basis in these roles.

David Hamilton et al., *Social Cognition and the Study of Stereotyping*, in *SOCIAL COGNITION: IMPACT ON SOCIAL PSYCHOLOGY* 291, 310 (Patricia G. Devine et al. eds., 1994).

110. Strong evidence of systematic and wide-spread manipulation of stereotypes in news reporting and other aspects of the mass media has been well-documented by FAIR (Fairness & Accuracy In Reporting), a national media watch group, in *Focus on Racism in the Media*, 5 EXTRA!, July/August 1992.

111. Devine, *supra* note 32, at 6.

112. See *supra* text accompanying notes 42-52.

113. See *supra* text accompanying note 51.

114. Ronis et al., *supra* note 98, at 232.



son's racial identity (or its symbolic equivalent). This means that for a person who rejects the stereotype to avoid stereotype-congruent responses to blacks (i.e., to avoid falling into a bad habit), she must intentionally inhibit the automatically activated stereotype and activate her newer personal belief structure. As Devine points out, "[s]uch inhibition and initiation of new responses involves controlled processes."<sup>115</sup> That is, "nonprejudiced responses take intention, attention, and effort."<sup>116</sup>

A particularly illuminating implication of this model is that unless a low-prejudiced person consciously monitors and inhibits the activation of a stereotype in the presence of a member (or symbolic equivalent) of a stereotyped group, she may unintentionally fall into the discrimination habit. For example, the whites in Duncan's study who interpreted the same ambiguous sleeve as hostile when the actor was black and as innocuous when the actor was white,<sup>117</sup> could have had well-internalized nonprejudiced beliefs. However, they may not have consciously monitored the automatic activation of the black stereotype. Because blacks are stereotypically viewed as hostile, activation of the stereotype would have primed the hostility category, making it more accessible for social judgments about the black actor. Since the black stereotype is automatically activated, it could have biased subjects' judgment of the black actor unconsciously.

One strength of this model, then, is that it explains how even people with well-internalized nonprejudiced standards are capable of unconscious discrimination against blacks. Patricia Devine designed an experiment whose results provide empirical support for this model. This research examined how automatic processes affected responses to members of a stereotyped group. The experiment involved presenting stereotype-related information to persons below their perceptual threshold, so that subjects could not consciously process the information. Thus, any effects of such subliminally presented information on subsequent social judgments would necessarily result from automatic processes. As discussed below, Devine found that the effects of automatic stereotype activation are equally strong and inescapable for high- and low-prejudiced subjects.

In Devine's study, both high- and low-prejudiced subjects performed a task that exposed them to either a low concentration (20% of a 100-word list) or a high concentration (80% of a 100-word list) of black stereotype labels (e.g., afro, lazy, musical, athletic, poor, etc.) in a manner determined to be effectively outside their conscious awareness.<sup>118</sup> For example, to prevent subjects from having conscious access to the labels, the labels were

---

115. Devine, *supra* note 32, at 6.

116. *Id.* at 16.

117. *See supra* note 90 and accompanying text.

118. Devine, *supra* note 32, at 9-10. To prevent subjects from having conscious access to the labels, or "primes," Devine presented the primes to the subjects' parafoveal visual field (i.e., outside the fovea, the area of the most distinct vision on the retina). A separate test with different subjects had determined that under these conditions subjects could not recall or recognize the primes. *Id.* at 8-10.

presented very rapidly (within a time frame of 80 milliseconds) and were followed immediately by a mask (i.e., a series of jumbled letters). None of these labels, or "primes," was related to hostility.<sup>119</sup> In an ostensibly unrelated second experiment, subjects read a behavioral description of a person named Donald, whose race was not specified, and who was engaging in a series of ambiguously hostile behaviors. For example, Donald demands his money back from a store clerk immediately after a purchase and refuses to pay his rent until his apartment is repainted.<sup>120</sup> Devine found that both high- and low-prejudiced subjects' ratings of the target's hostility were significantly higher (i.e., indicated more hostility) when subliminally exposed to a high, rather than a low, concentration of black-stereotype labels.<sup>121</sup>

These findings demonstrate that well-learned sets of associations like stereotypes can be activated automatically in perceivers' memories and can affect subsequent social judgments. The effects of automatic stereotype priming on subjects' evaluation of the target person's hostility are especially revealing in Devine's experiment because no hostility-related traits were used as primes. Thus, it seems that the black stereotype must be constructed cognitively in such a way that activating one component of the stereotype simultaneously primes or activates the remaining closely associated components as well.<sup>122</sup> These findings also suggest that even low-

---

119. One set of primes included the following twelve words: nigger, poor, afro, jazz, slavery, musical, Harlem, busing, minority, oppressed, athletic, and prejudice. The other set of primes included the following: Negroes, lazy, Blacks, blues, rhythm, Africa, stereotype, ghetto, welfare, basketball, unemployed, and plantation. *Id.* at 10.

120. *Id.*

121. *Id.* at 11-12.

122. Research shows that the accessibility of a construct will also increase if the accessibility of a closely related category is increased. For example, studies in cognitive literature have shown that subjects recognize or process a word faster when the word (e.g., butter) is preceded by an associated word (e.g., bread) than an unassociated word (e.g., nurse). See generally David E. Meyer et al., *Loci of Contextual Effects on Visual Word Recognition*, in ATTENTION AND PERFORMANCE 98 (Patrick M.A. Rabbit & Stanislaw Dornic eds., 1975). Research on social cognition has found the same priming effects for closely associated social constructs. For example, a 1955 study found that activating one social construct (e.g., reading about the life of Pope Pius XII) will increase the accessibility of closely related social constructs (e.g., the principles of the Catholic Church). Higgins & King, *supra* note 82, at 82-83. There is also evidence that constructs with a similar evaluative tone are closely related, so that activating a social construct (e.g., ugly) should also increase the accessibility of evaluatively similar social constructs (e.g., evil, gangster, mugging). See Delos D. Wickens, *Characteristics of Word Encoding*, in CODING PROCESSES IN HUMAN MEMORY 191 (Arthur W. Melton & Edwin Martin eds., 1972).

One useful framework developed by cognitive psychologists for explaining the resonance of frequently associated categories is the associative network model. See Thomas Ostrom et al., *The Cognitive Foundation of Attitudes: It's a Wonderful Construct*, in SOCIAL COGNITION: IMPACT ON SOCIAL PSYCHOLOGY 195, 212, 224 (Patricia G. Devine et al. eds., 1994). Network models assume that information is stored as *nodes* in memory, with *pathways* connecting to the nodes. *Id.* (emphasis added). The nodes are thought to represent different types of information, including the full complement of constructs that constitute the world as each of us experiences it, while the pathways represent "the associations formed between the nodes." *Id.* at 224. When a node is activated, the network models postulate that excitation "radiates" along the pathways to other closely associated nodes. *Id.* The effect of this spreading excitation is to prime or activate the nodes linked to it. Moreover, some pathways are more efficient conductors of excitations than others, with the efficiency of conduction thought to be "a

prejudiced subjects who have well-internalized nonprejudiced beliefs about blacks have cognitive structures (i.e., stereotypes) that automatically produce stereotype-congruent evaluations of ambiguous behaviors when subjects cannot monitor stereotype activation consciously.

In this sense, Lawrence and other commentators are correct in pointing out that we are all prone to stereotype-congruent or prejudice-like responses to blacks and other stereotyped groups. However, the research demonstrating disassociation of stereotypes and personal beliefs in low-prejudiced people argues against the conclusion that "we are all racists." Instead, I suggest that it is more accurate and useful to say that "we are all creatures of habit."

### III

#### COMBATting UNCONSCIOUS DISCRIMINATION IN THE COURTROOM

Perhaps the most important strength of the dissociation model of automatic and controlled processes I have outlined is that it suggests a strategy for resisting unconscious discrimination. Thus far, the model has focused on how people who are firmly committed to their low-prejudiced beliefs<sup>123</sup> remain prone to automatic activation of stereotypes. According to the model, for such individuals to resist falling into the discrimination habit, they repeatedly recall their personal beliefs so that their social judgments become based on these beliefs rather than the stereotypes. Reminding decisionmakers of their personal beliefs, therefore, may help them to resist fall-

---

direct function of the frequency with which two constructs have been associated." *Id.* Thus, for example, if the constructs *money* and *bank* frequently have been associated in a person's life, there will be an efficient pathway linking them; if not, the pathway will be less efficient. The more efficient the pathway linking two constructs—i.e., the more frequently two constructs have been associated in a person's life—the more readily will activation of one construct simultaneously prime or activate its associate. *Id.*

123. Although the dissociation model focuses on individual-level conflict between more egalitarian and less prejudiced beliefs on the one hand, and persistent prejudice-like responses on the other, I am not suggesting that high-prejudiced people no longer exist. Jennifer L. Hochschild and Monica Herk, for example, have pointed out that a small group of "hard-core" racists still exists in America. Jennifer L. Hochschild & Monica Herk, "Yes, but . . .": *Principles and Caveats in American Racial Attitudes*, in NOMOS XXXII: MAJORITIES AND MINORITIES 308, 311 (John W. Chapman & Alan Wertheimer eds., 1990). As recently as 1978, as many as 15% of white adults believed that "blacks are inferior to white people." *Id.* (Moreover, with the publicity surrounding recently released books like *The Bell Curve*, which is essentially an apology for the subordination of blacks in this country, the number of people who endorse derogatory black stereotypes may be on the rise.) Reducing the bias of "hard-core" racists—people whose personal beliefs overlap derogatory black stereotypes—poses special challenges for lawyers and judges concerned about reducing courtroom bias. Perhaps Rokeach's self-confrontation technique, see *infra* note 142, could be employed in some extreme cases. However, in light of social science data indicating that prejudice has been declining steadily over the past 40 years, see *supra* text accompanying notes 20-21, I am proceeding on the assumption that many (if not most) present-day jurors hold nonprejudiced personal beliefs that may be activated in the service of less biased social judgments. Insofar as this assumption does not hold in a given case, other techniques (e.g., Rokeach's self-confrontation technique) may be called for.

ing unconsciously into the discrimination habit. This section describes evidence that this approach can be effective.

### A. Empirical Support for the Strategy

The range of discrimination-reduction strategies that courts will countenance and lawyers will employ may turn significantly on which model of prejudice they adopt. Whereas the unconscious-racism model fails to explore the utility of conscious processes and nonprejudiced beliefs in the fight against racial discrimination, the present model views such processes and beliefs as valuable weapons in the fight against discrimination.

There is considerable empirical evidence supporting the dissociation model's assumption that responses based on automatic processes can be inhibited and replaced by responses based on controlled processes.<sup>124</sup> Focusing on the effect of gender stereotypes on memory, for example, Higgins and King demonstrated that when gender was not brought situationally to subjects' attention, or made "salient,"<sup>125</sup> subjects' descriptions of self and others reflected more traditional views of gender-linked attributes.<sup>126</sup> Higgins and King have suggested that, under such conditions,

---

124. For example, by using a semantic priming task in which the presentation of a word unconsciously influences subsequent processing of related words, James H. Neely has demonstrated that when automatic processing produces a response that conflicts with conscious expectation (induced through experimenter instructions), subjects inhibit the automatic response and intentionally replace it with one consistent with their conscious expectation. James H. Neely, *Semantic Priming and Retrieval from Lexical Memory: Roles of Inhibitionless Spreading Activation and Limited-Capacity Attention*, 106 J. EXPERIMENTAL PSYCHOL. 226, 251-53 (1977).

125. The salience of the subjects' gender was manipulated by varying the sexual composition of the different experimental groups. For example, in one series of experiments, a female experimenter conducted 20 groups of subjects each composed of two or three females and one male, and a male experimenter conducted 20 groups of subjects each composed of two or three males and one female. Higgins & King, *supra* note 82, at 86. The researchers reasoned that "gender should be more salient for a group member whose gender is in the minority than for a group member whose gender is in the majority." *Id.* at 85.

126. *Id.* at 87-88. In Study 1, the subjects (146 college students) in half of the groups read a paragraph supposedly describing a female undergraduate at Princeton (Barbara) and subjects in the other half of the groups read the same paragraph ambiguously describing a male undergraduate (Bob). *Id.* at 86-87. The paragraph was constructed to unambiguously exemplify the following eight traits: two evaluatively positive, stereotypically male traits, (active, ambitious) and two evaluatively negative, stereotypically male traits, (aggressive, selfish); two evaluatively positive, stereotypically female traits, (polite, sensitive) and two evaluatively negative, stereotypically female traits, (emotional, dependent). *Id.* at 86. After reading the paragraph, the subjects were given a 20-minute filler task to take their minds off the paragraph they had just read. They were then asked to reproduce the paragraph about Bob (Barbara) as best they could, word for word. *Id.* at 87.

When the person described in the paragraph was ostensibly male (Bob), the subjects recalled less stereotypically male and more stereotypically female information about him when their gender was in the minority (i.e., high gender salience) than when their gender was in the majority (i.e., low gender salience). *Id.* Moreover, when the person described in the paragraph was ostensibly female (Barbara), the subjects recalled less stereotypically female information about her when their gender was in the minority than when it was in the majority. *Id.* Similar results were found in a different study in which subjects (58 college students) were again put into groups containing a solitary male or female and simply asked to "tell us about yourself" in writing. *Id.* at 103. "[M]ales in the minority described themselves as more stereotypically female and less stereotypically male than males in the majority,

traditional gender stereotypes, with their longer history and greater frequency of activation, are activated automatically and influence recall.<sup>127</sup> When gender was brought to the subjects' attention or made salient, however, they apparently inhibited the traditional stereotype, and descriptions were more consistent with their more recently developed, modern views of gender-linked attributes.<sup>128</sup> In other words, when the subjects were reminded of gender, they checked their stereotype-congruent responses more assiduously than when gender was less salient.

This dissociation model is also consistent with the findings, described earlier, concerning the proclivity of white rescuers to help white but not black victims in distress. Recall that white research subjects led to believe that a person was in distress helped black victims as often as white victims when there was no ostensible justification for a failure to help.<sup>129</sup> On the other hand, if the subjects knew of the availability of another who might help, they "helped black victims much less frequently than they helped white victims."<sup>130</sup> According to the dissociation model, when the subjects believed they were the only potential rescuer, they were required consciously to think about what their responses to the black victim's call of distress implied about their nonprejudiced self-conceptions. When the conflict between their nonprejudiced personal beliefs and the stereotype of blacks is made salient in this way, the dissociation model predicts that low-prejudiced persons are likely to resolve the conflict by inhibiting their prejudice-like responses and reaffirming their nonprejudiced self-conceptions. On the other hand, when the subjects believed that there were others who might help, the stereotype-personal belief conflict was less salient and the low-prejudiced subjects were therefore less likely to monitor and inhibit responses based on the negative black stereotype.<sup>131</sup>

---

whereas females in the minority described themselves as more stereotypically male and less stereotypically female than females in the majority." *Id.* at 104. Thus, when the college students' gender was relatively salient, they were more likely to describe themselves in terms of nontraditional or modern views of male and female attributes. *Id.* Otherwise, their spontaneous self-descriptions (as well as their descriptions of others) tended to reflect the prevailing sexual stereotypes. *Id.*

127. *Id.* at 85-86.

128. *Id.* at 87-88, 104-05; *cf. supra* text accompanying notes 53-56 (suggesting similar psychological processes involved in respondents' expressions of nonprejudiced thoughts about blacks in an anonymous thought-listing task).

129. *See supra* text accompanying notes 68-70.

130. Gaertner & Dovidio, *supra* note 61, at 77.

131. *See* Monteith, *Self-Regulation*, *supra* note 52, at 478-84 (describing an experiment in which low-prejudiced subjects effectively inhibited prejudiced responses to jokes about gays when experimenters alerted them to discrepancies between their nonprejudiced personal standards and their discriminatory tendencies).

*B. A Narrative of Hope*

"YOU'RE ALL PREJUDICED" DARROW TELLS JURY

[Headline, *Detroit Free Press*, May 12, 1926]<sup>132</sup>

Clarence Darrow, one of the most famous lawyers in American history, directly challenged jurors to confront their own prejudices in a dramatic murder trial early in this century. The case involved a black family who moved into a middle-class white Detroit neighborhood in 1925. When Dr. Ossian Sweet and his wife moved into the neighborhood with their baby daughter, they knew other blacks who had bought homes in white neighborhoods had been forced to move by "Improvement Associations."<sup>133</sup> Accordingly, Dr. Sweet brought along his brothers, several friends, and an ample supply of guns and ammunition. Two nights after his arrival, a large white crowd, estimated at several hundred, gathered around the house and began throwing stones at the house amid cries of "Niggers."<sup>134</sup> Although police officers were present to maintain order, they stood idly by as the barrage of rocks increased. Seeing a big stone crash through an upstairs window and the crowd make a sudden movement, both Sweet and his younger brother fired a warning shot over the heads of the boisterous mob. One of the mob's members was killed.<sup>135</sup>

Everyone in the house was arrested and charged with murder. The NAACP asked Darrow to come out of retirement to defend the Sweets. Darrow agreed.<sup>136</sup> In his summation to the jury, Darrow challenged them to confront their own racial biases directly:

I haven't any doubt but that every one of you is prejudiced against colored people. I want you to guard against it. I want you to do all you can to be fair in this case, and I believe you will. . . .

. . . .

You need not tell me you are not prejudiced. I know better. We are not very much but a bundle of prejudices anyhow. We are prejudiced against other people's color. Prejudiced against other men's religions; prejudiced against other people's politics. Prejudiced against people's looks. Prejudiced about the way they dress. We are full of prejudices. . . .

. . . .

. . . Here were eleven colored men, penned up in the house. Put yourselves in their place. Make yourselves colored for a little while. It won't hurt, you can wash it off. They can't, but you can; just make yourself black for a little while; long enough, gentlemen, to

---

132. Arthur Weinberg, *You Can't Live There!*, in *ATTORNEY FOR THE DAMNED* 229, 229 (Arthur Weinberg ed., 1957).

133. *Id.* at 230.

134. *Id.*

135. *Id.*

136. *Id.*

judge them, and before any of you would want to be judged, you would want your juror to put himself in your place. That is all I ask in this case, gentlemen. They were black, and they knew the history of the black. . . .

. . . .  
. . . Supposing you had your choice, right here this minute, would you rather lose your eyesight or become colored? Would you rather lose your hearing or be a Negro? Would you rather go out there on the street and have your leg cut off by a streetcar, or have a black skin? . . .

. . . Life is a hard game anyhow. But, when the cards are stacked against you, it is terribly hard. And they are stacked against a race for no reason but that they are black.<sup>137</sup>

The jury returned a not guilty verdict for Dr. Sweet, and the prosecution decided not to proceed further against any of the remaining defendants.<sup>138</sup>

Dr. Sweet's case provides a compelling narrative of hope and redemption that stands in marked contrast to the pessimism of many current discussions of prejudice in the courtroom. Clarence Darrow, in the heyday of Jim Crow, successfully urged a jury of all white males to resist succumbing to their discriminatory impulses in judging the reasonableness of a black man's use of lethal force against a white man. Darrow's feat was especially remarkable because it required Darrow to combat the influence of both stereotypes and prejudice on the factfinders. In the 1920s, just as today, American culture was replete with derogatory images of blacks. Thus, negative black stereotypes that could be triggered automatically by the presence of a black person were well established in the factfinders' memories. Moreover, the percentage of whites who accepted or endorsed the prevailing black stereotypes was much greater in the past than it is today.<sup>139</sup> Many of Dr. Sweet's jurors, therefore, probably also formed a conscious expectation for instances of trait categories stereotypically associated with blacks (e.g. D.W. Griffith's popular and celebrated 1915 film, *Birth of a Nation*, presented a Ku Klux Klan view of blacks as lawless savages). Because automatic (stereotype-driven) and controlled (prejudice-driven) processes can operate simultaneously on the same underlying categories,<sup>140</sup> the two

---

137. *Id.* at 234-35, 252, 257 (quoting Clarence Darrow's closing argument).

138. *Id.* at 263.

139. See *supra* text accompanying notes 20-21.

140. Recall Harold Kelley's study of labeling effects on impression formation discussed *supra* at note 95. In that study, students' ratings of a new instructor were more favorable when the experimenter described the instructor as a "warm" person to the students before the instructor's arrival than when the experimenter described the instructor as a "cold" person. Kelley's interpretation of these results was that the experimenter's prior description gave the students a conscious expectation for instances of the category designated by the label (e.g. "warm") as well as other categories such as "friendly" and "helpful" assumed to be closely related to that category. That is, the students' ratings were driven by controlled processes. Kelley, *supra* note 95, at 433-35. Higgins and King suggest that another possibility is that exposure to a particular label automatically activated the category designated by the

processes likely were mutually reinforcing in many of these jurors; that is, both processes combined additively to make the underlying negative categories about blacks more accessible.<sup>141</sup> Confronting factfinders whose personal beliefs and stereotypes about blacks overlapped, Darrow's strategy was based on the assumption that even high-prejudiced persons personally endorse general egalitarian beliefs. Dr. Sweet's life hinged on whether the jurors—prompted by Darrow's race-conscious appeals—could resist their discriminatory impulses and respond to Dr. Sweet on the basis of their egalitarian ideals.<sup>142</sup> Fortunately for Dr. Sweet and those of us who find relief from despair in what his case says about the capacity of jurors to resist even their most entrenched biases, the jury responded to Darrow's plea by activating their egalitarian responses and checking their prejudiced and stereotype-congruent ones.

Today, although more people espouse nonprejudiced personal beliefs than in Darrow's time, the black stereotype is probably no less entrenched in the memories of Americans than in Darrow's day. Indeed, with the advent of an omnipresent mass media and its incessant manipulation of stereotypes,<sup>143</sup> it may be more entrenched. Thus, habitual stereotype-congruent responses to blacks, even by sincerely racially liberal whites, may distort legal judgments concerning blacks as much in contemporary America as in the America Darrow knew. If so, these distorted judgments are more insidious than before because they result from automatic processes, which often (but not necessarily always) escape conscious detection. Nevertheless, Darrow's strategy of explicitly engaging our egalitarian responses and urging us consciously to substitute them for our more habitual responses squares with modern empirical research on discrimination

---

label as well as other categories assumed to be closely related to that category. From this perspective, the students' ratings were driven by automatic processes. Higgins & King, *supra* note 82, at 73-74. In the end, Higgins and King suggest that the students' ratings were probably influenced by both controlled and automatic processes; in their words, Kelley's "findings are likely to have been multiply determined." *Id.* at 74. See also John A. Bargh et al., *The Additive Nature of Chronic and Temporary Sources of Construct Accessibility*, 50 J. PERSONALITY & SOC. PSYCHOL. 869, 870 (1986) (discussing how different accessibility influences operating within the same person at the same time "may be mutually facilitative, such that both serve to make a single construct more accessible").

141. See *id.* at 870, 876-77.

142. Darrow was employing a strategy for changing highly prejudiced attitudes similar to what fifty years later came to be known as the Rokeach's confrontation technique. See MILTON ROKEACH, *THE NATURE OF HUMAN VALUES* 286-306 (1973). Rokeach's strategy was inspired by Gunnar Myrdal's characterization of the American Dilemma. 1 GUNNAR MYRDAL, *AN AMERICAN DILEMMA: THE NEGRO PROBLEM AND MODERN DEMOCRACY* 89, 1139-41 (1944). According to Myrdal, many white Americans, who are committed to the general egalitarian tenets of the "American Creed," but who simultaneously have specific prejudiced tendencies, experience an internal moral conflict. Rokeach reasoned that self-dissatisfaction arises when people are encouraged to recognize the discrepancy between their egalitarian self-conception and their prejudiced responses. This self-dissatisfaction should then motivate individuals to change the prejudiced aspects of their responses to be more in line with their egalitarian self-image. ROKEACH, *supra*, at 286.

143. See *supra* note 110.



reduction techniques.<sup>144</sup> As the recent research in social cognition demonstrates, avoiding stereotype-congruent responses requires conscious effort by the decisionmaker.<sup>145</sup>

A more recent trial, the infamous New York subway vigilante case of *People v. Goetz*,<sup>146</sup> is an example of a case in which the failure to give direct and explicit consideration to racial factors may have resulted in less fair deliberations by the factfinders. In this case, the defendant, Bernhard Goetz, successfully claimed that his shooting of four black teenagers after one of them requested five dollars was justified as an act of self-defense.<sup>147</sup> Professor George Fletcher, a legal theorist who witnessed the entire trial, identified numerous unmistakable instances of the defense "indirectly and covertly . . . play[ing] on the racial factor."<sup>148</sup> For example, the defense "relentlessly attack[ed]" the black victims as "savages," "vultures," the "predators" on society, and the "gang of four."<sup>149</sup> According to the dissociation model developed in Part II, the use of such racial imagery<sup>150</sup> automatically activates and reactivates the black stereotype. This activation renders negative thoughts and *feelings*<sup>151</sup> associated with that stereotype acutely accessible for social judgments about the black victims. The factfinders

144. See Monteith, *Self-Regulation*, *supra* note 52, at 478-84.

145. See *supra* text accompanying notes 53-56, 124-128.

146. 497 N.E.2d 41 (N.Y. 1986).

147. *Id.* at 43-46. However, the court reinstated the dismissed counts of the indictment on other grounds. *Id.* at 53-54.

148. GEORGE P. FLETCHER, A CRIME OF SELF-DEFENSE: BERNHARD GOETZ AND THE LAW ON TRIAL 206 (1988). One such trial tactic involved recreating the shooting of the teenagers, for which the defense called in four "props" to act as the four black victims:

The nominal purpose of the demonstration was to show the way in which each bullet entered the body of each victim. The defense's real purpose, however, was to re-create for the jury, as dramatically as possible, the scene that Goetz encountered when four young black passengers began to surround him. For that reason [Goetz's attorney] asked the Guardian Angels to send him four young black men to act as the props in the demonstration. In came the four young black Guardian Angels, fit and muscular, dressed in T-shirts, to play the parts of the four victims in a courtroom minidrama.

*Id.* at 207. Although the witness who these black men surrounded was not authorized to testify about the typical person's fear of being accosted by four such individuals, the defense "designed the dramatic scene so that the implicit message of menace and fear would be so strong that testimony would not be needed." *Id.* at 130.

149. *Id.* at 206. As Fletcher insightfully notes:

These verbal attacks signaled a perception of the four youths as representing something more than four individuals committing an act of aggression against a defendant. That "something more" requires extrapolation from their characteristics to the class of individuals for which they stand. There is no doubt that one of the characteristics that figures in this implicit extrapolation is their blackness.

*Id.*

150. "[A]ny word, metaphor, argument, comment, action, gesture, or intonation" that resonates with derogatory racial stereotype constitutes racial imagery. Johnson, *supra* note 16, at 1799.

151. To this point, I have focused on the descriptive content of stereotypes. Recent research has also begun to investigate the *automatic affective* processing effects associated with racial category priming. Negative attitudes and emotional responses to blacks—hatred, loathing, disgust, fear, etc.—have historically accompanied the negative personality traits associated with blacks. Thus, priming the racial category probably also primes negative responses that are associated with the category. This proposition is consistent with Gordon Allport's observation that a category such as a racial group

may not experience these judgments as stemming from their knowledge of the black stereotype, but instead as rational, evenhanded evaluations of objective reality. The only way to combat the effect of covert appeals to racial imagery may be to challenge explicitly the factfinders to monitor consciously their responses to avoid unconscious stereotyping. In considering the impact on the jury of the defense's use of racial imagery, Fletcher reaches conclusions about the need to openly address the racial factor that are consistent with this analysis:

In the end, Slotnick's [Goetz's attorney] covert appeal to racial fear may have had more impact on the jury precisely because it remained hidden behind innuendo and suggestion. It spoke to that side of the jurors' personality that they could not confront directly. Paradoxically, Slotnick may have gained more from not [explicitly playing on the racial factor] than from bringing the racial issue out into the open. Openly talking about racial fear in the courtroom might have helped the jury to deal more rationally with their own racial biases.<sup>152</sup>

### *C. A Blueprint for Applying the Dissociation Model*

References to stereotyped groups in legal proceedings vary in both content and subtlety. The content of a group reference concerns the specific aspect of the stereotype that the reference invokes. For example, fairly recent cases record attorneys playing on stereotypes of blacks,<sup>153</sup> Italians,<sup>154</sup> and Native Americans<sup>155</sup> as more prone to violence and criminality than other Americans. Other cases describe attorneys invoking stereotypes of blacks as subhuman, sexually predatory, and dishonest, to name a few.<sup>156</sup> As the earlier analysis revealed, playing on any particular aspect of a stereotype may activate the entire stereotype,<sup>157</sup> thus distorting a wide range of social judgments about the stereotyped litigant.

---

category "saturates all that it contains with the same ideational and emotional flavor." ALLPORT, *supra* note 35, at 21.

152. FLETCHER, *supra* note 148, at 208.

153. *Smith v. State*, 516 N.E.2d 1055, 1064 (Ind. 1987), *cert. denied*, 488 U.S. 934 (1988) (prosecutor's statement that defendant "had to play Superfly," a fictional black criminal popularized in motion pictures); *People v. Lurry*, 395 N.E.2d 1234, 1237 (Ill. App. Ct. 1979) (prosecutor's lengthy statement about prevalence of black crime, urging the jury to prevent this pattern of violence from reaching their city); *State v. Noel*, 693 S.W.2d 317, 318 (Mo. Ct. App. 1985) (prosecutor's statement that "[n]inety percent of all murders are committed by blacks on blacks" and that "[i]ts [sic] time to say 'We're not going to allow this kind of conduct to go on in our city anymore.'").

154. *Commonwealth v. Graziano*, 331 N.E.2d 808, 812 (Mass. 1975) (in a case involving recent immigrants, prosecutor's reference to Al Capone and "the godfather").

155. *Soap v. Carter*, 632 F.2d 872, 878 (10th Cir. 1980) (Seymour, J., dissenting), *cert. denied*, 451 U.S. 939 (1981) (in case involving American Indians, prosecutor's statement that "when you see an Indian that drinks liquor, you see a man that can't handle it" and that "they can't manage it").

156. See Johnson, *supra* note 16, at 1753-56.

157. See *supra* text accompanying note 122.

The subtlety of group references ranges from blatant to covert and indirect. Although recent case law contains numerous examples of references from both ends of this continuum,<sup>158</sup> the dissociation model suggests that the more subtle references may be particularly pernicious because subtle references inconspicuously may activate the relevant stereotypes.<sup>159</sup> This deprives jurors of the opportunity to monitor consciously their responses for convergence with their personal beliefs.<sup>160</sup> Moreover, for courts concerned with policing inappropriate group references, subtle references pose special problems of identification. For example, Goetz's attorney Barry Slotnick did not explicitly mention race in his characterization of the black youths his client shot as "savages" and "predators," but his statements arguably constituted subtle racial references.

To identify the covertly racial tenor of such statements, courts need a test of the symbolic significance that the culture attaches to them. For to the extent that certain references carry racial connotations, they constitute symbolic equivalents of members of that race and thus serve as cues that activate (often unconsciously) racial stereotypes. Thoughtful formulations of tests for identifying subtle racial symbolism have been developed by Professor Lawrence (the cultural meaning test)<sup>161</sup> and Professor Johnson (the racial imagery shield law).<sup>162</sup> Whatever test for identifying references

158. See Johnson, *supra* note 16, at 1760-62.

159. See *supra* text accompanying notes 118-122.

160. See *supra* text accompanying notes 124-130.

161. See Lawrence, *supra* note 13, at 355 (arguing that the prevalence of unconscious discrimination thwarts intent analysis). Lawrence argues for a "cultural meaning test," by which governmental conduct would be evaluated "to see if it conveys a symbolic message to which the culture attaches racial significance." *Id.* at 356. See also Elizabeth L. Earle, *Banishing the Thirteenth Juror: An Approach to the Identification of Prosecutorial Racism*, 92 COLUM. L. REV. 1212, 1239-40 (1992) (arguing that adapted to prosecutorial forensics, Lawrence's cultural meaning test can become a reasonable person standard—i.e., whether a reasonable person would construe the reference to have a racial tenor—and that indirect references to race should be gauged according to this standard).

162. Johnson, *supra* note 16, at 1799-1800. Employing the term "racial imagery" to refer to any racial reference, whether blatant or subtle, Johnson proposes the following test for racial imagery:

"Racial imagery" is any word, metaphor, argument, comment, action, gesture, or intonation that suggests, either explicitly or through commonly understood allusion, that

(1) a person's race or ethnicity affects his or her standing as a full, capable, and decent human being; or

(2) a person's race or ethnicity in any way affects the credibility of that person's assertions; or

(3) a person's race or ethnicity in any way affects the likelihood that he or she would choose a particular course of conduct whether criminal or noncriminal; or

(4) a person's race or ethnicity in any way affects the appropriate sanctions for a crime committed by or against him or her; or

(5) a person's race or ethnicity sets him or her apart from members of the jury, or makes him or her allied with members of the jury or, more generally, that a person's race or ethnicity allies him or her with other persons of the same race or ethnic group or separates him or her from persons of another race or ethnic group.

Racial imagery will be conclusively presumed from the unnecessary use of a racially descriptive word.

Where a metaphor or simile uses the words "white," "black," "brown," "yellow," or "red"; where any comparisons to animals of any kind are made; or where characters, real or fictional, who are strongly identified with a racial or ethnic group are referred to, racial

a court adopts, fairness and accurate factfinding require that once the court identifies an inappropriate reference, it should give the opposing party the choice of a mistrial or corrective instructions. Given the enormous societal interest in racially fair legal proceedings, courts must follow a policy of "zero tolerance" with respect to inappropriate racial references.

### 1. *Admission of Rationality-Enhancing Group References*

The central thesis of this article, however, is that not all references to blacks or other stereotyped groups are inappropriate. I therefore turn now to the critical distinction between rationality-enhancing and rationality-subverting group references. Group references that exploit, exacerbate, or play on the prevailing stereotypes that factfinders carry with them into the jury box subvert the rationality of the fact-finding process. But references that challenge the factfinders to reexamine and resist their discriminatory responses enhance the rationality of the fact-finding process. For example, referring to blacks as animal-like or subhuman (e.g., "savages," "monsters," and "Tasmanian devils"<sup>163</sup>) resonates so strongly with prevailing black stereotypes as to constitute rationality-subverting racial references. In contrast, Darrow's plea for jurors consciously to monitor and resist their anti-black prejudices and stereotypes is an example of rationality-enhancing uses of group references.

I propose that courts recognize and apply this distinction as follows. Once a court identifies a group reference made by one litigant's counsel,<sup>164</sup> it should give the opposing side the choice of a mistrial or corrective instructions. This general proscription of group references should be subject to

---

imagery will be presumed, subject only to rebuttal through proof that the term in question could not have racial connotations with respect to any witness, defendant, attorney, or judge involved in the case.

That a speaker disclaims racial intent, either contemporaneously or at a later date, shall have no bearing upon the determination of whether his or her remarks or actions constitute a use of racial imagery.

*Id.*

163. In the first trial of the four white Los Angeles police officers who beat Rodney King, one of the defendants, Stacey C. Koon, testified that King was "a monster-like figure akin to a Tasmanian devil." *Latest Defense Witness in Rodney King Trial Backfires*, L.A. SENTINEL, Apr. 1, 1993, at A4.

164. A group reference should be chargeable to an attorney if he knows or reasonably should know that a witness will make such a reference and does not take reasonable steps (e.g., remonstrance of the witness) to prevent the reference. Under the current code of professional conduct, an attorney is obligated not to offer false evidence, including putting a witness on the stand when she anticipates false testimony. MODEL RULES OF PROFESSIONAL CONDUCT, Rule 3.3(a)(4) (1994). Our growing understanding of social cognition suggests that inappropriate and unchecked references to stereotyped groups (including their symbolic equivalent) subvert justice and accurate truthseeking as much as perjury. In fact, covert group references may subvert truthseeking even more than perjury to the extent that jurors discount the testimony of defendants, based on their common sense understanding that such witnesses have substantial incentives to lie. But covert group references can distort jurors' social judgments unconsciously, thus depriving them of the opportunity to weigh and discount the reference in the way they could the testimony of defendants.

very limited exceptions.<sup>165</sup> One such exception would be that the reference enhances rationality, in that it challenges factfinders to monitor and inhibit their stereotype-congruent responses.<sup>166</sup> To fall under this exception, the attorney making or seeking to make the reference in litigation would have to represent the interests of the member of that group and hold a good faith belief in the rationality-enhancing value of the reference. Inasmuch as courts seek to protect the truthseeking function of the trial process, I argue that they should maintain the same goal in cases involving litigants from other stereotyped groups.

Especially sensitive are those cases in which an attorney representing a member of a stereotyped group makes an ostensibly derogatory reference to that group. The attorney may claim that such a reference actually helps factfinders resist prejudice-like responses when making judgments about the group member whom she represents. For example, on one level Darrow's statement that the white jurors would rather lose their eyesight and have their legs cut off by a streetcar than have black skin makes a very invidious statement about the negative social value of his client's racial identity. (Today white college students regularly report that if they were suddenly to become outwardly black while they inwardly remain who they were, reasonable compensation would be one million dollars a year for life!<sup>167</sup>) On another level, however, it was clear from the context in which Darrow used these ostensibly invidious racial references that they actually constituted an appeal to the jurors to rise above their prejudices on that occasion.

Similarly, in the movie *Philadelphia*, the attorney's use of the epithets associated with gay men—"faggot," "fairy," "booty snatcher," "rump-roaster"—were calculated to challenge the factfinders to confront and inhibit their habitual stereotype-congruent responses to gays, not to encourage such responses. Not mentioning race or sexual orientation at all, when cues that automatically trigger the stereotypes associated with these characteristics abound, promotes unconscious discrimination against members of these stereotyped groups. To control a bad habit, a person first must

---

165. An example of a proper reference to race would be one in which race is part of a description given for the purpose of identification.

166. In arguing that colorblind legal proceedings may impair the capacity of legal decisionmakers to inhibit habitual prejudice-like responses and activate newer nonprejudiced ones, I decidedly am not suggesting that legal proceedings should be open to any and all racial references, however invidious. To the contrary, in *Armour*, *supra* note 16, I argued unequivocally against the use of race-based claims of reasonableness in self-defense claims, even though race may be formally relevant to such claims. On a strictly formal level, therefore, it may seem that I am taking a paradoxical position on the appropriateness of racial references in legal proceedings: legal decisionmakers should and should not consider race in resolving legal disputes. The paradox dissolves, however, as soon as we scratch beneath the superficial layer of formalism and consider the distinction between racial references that subvert the rationality of the fact-finding process and racial references that enhance that rationality. *Id.*

167. ANDREW HACKER, *TWO NATIONS* 32 (1992).

recall it consciously, and then intentionally inhibit it as he or she responds in ways consistent with his or her personally endorsed beliefs and attitudes.

## 2. *Timing of Rationality-Enhancing Group References*

Assuming that an attorney representing a member of a stereotyped group seeks to make rationality-enhancing group references for her client, what is the best setting for such references? Salutory group references may be made in a variety of settings, including voir dire, opening statements, presentation of the case in chief, closing arguments, and jury instructions. All these settings present opportunities for legal actors to encourage prospective or sitting factfinders to guard against their prejudice-like responses. *Voir dire*, for example, may present an excellent opportunity to search not only for avowedly prejudiced venirepersons, but also to signal to prospective jurors the importance of consciously monitoring their habitual responses to the stereotyped litigant.

Another favorable setting for rationality-enhancing group references may be during opening statements. A striking illustration of both a lawyer's efforts to activate a jury's nonprejudiced impulses in this setting, and a court's misguided application of color-blind formalism to his efforts, comes from a tort case pitting a pharmaceutical company against a low-income black infant and her mother.<sup>168</sup> The mother and child sought damages for serious injuries the child suffered from using the pharmaceutical company's allegedly defective drug. In his opening statement to a mostly white jury, plaintiffs' counsel characterized the case as "a test of our judicial system to see if a child who is at the lower end of our society . . . can come before a jury and receive fair and just compensation for [her] injuries."<sup>169</sup> He then directly addressed the racial dimension of the case:

[W]e were concerned about the effect of having black people come to an area where there are not many black people and expecting to get justice from a jury which is mostly white people. We decided to confront this issue and we asked you the questions this morning, and we were really pleased with the responses that we got and we think that this is an impartial jury and everyone here has sworn that they will try this case not on the basis of passions, or prejudice, or economic basis, but on the basis of the facts and the law.<sup>170</sup>

Criticizing counsel's comments, the Third Circuit warned that "the remarks should not be repeated in the opening statement at the retrial."<sup>171</sup>

---

168. *Stanton by Brooks v. Astra Pharmaceutical Products, Inc.*, 718 F.2d 553 (3d Cir. 1983). I am indebted to Professor Frank McClellan for bringing this case to my attention. Professor McClellan was co-counsel on this case. Both he and his partner are black, and they frequently try cases before predominantly white juries on behalf of black plaintiffs.

169. *Id.* at 578 (quoting from the trial transcript).

170. *Id.* at 578-79 (quoting from the trial transcript).

171. *Id.* at 579.

According to the court, the counsel's remarks were "beyond the realm of appropriate advocacy . . . '[T]here must be limits to pleas of pure passion and there must be restraints against blatant appeals to bias and prejudice.'"<sup>172</sup>

In this case, the Third Circuit elided the crucial distinction between rationality-enhancing and rationality-subverting racial references that I have been developing. A large and compelling body of social science research—including case studies, laboratory findings in mock jury studies, and general research on racial prejudice—establishes that racial bias affects jury deliberations.<sup>173</sup> Perhaps the best chance for litigants like this low-income black infant and her mother to receive a fair trial from a mostly white jury is for courts to recognize this distinction and permit rationality-enhancing racial references by a black litigant's attorney in settings likely to maximize the salutary effects of such references.

What settings are most favorable for rationality-enhancing group references? While this important question requires detailed consideration that is beyond the reach of this article, I suggest a few ideas to serve as a starting point for examining this issue. Confronting individuals about their biased, unconscious reflexes may produce different results at different stages of the litigation process. Jury studies suggest that many jurors already have made up their minds before the closing argument.<sup>174</sup> This suggests that perhaps *voir dire* and opening statements (and even pre-trial publicity) are the most important settings for rationality-enhancing comments and confrontations. In Darrow's celebrated defense of Dr. Sweet, for example, one wonders whether, in addition to his closing argument, he also confronted the jurors' stereotypes and prejudices at earlier stages of the proceeding, and if so, whether these earlier confrontations significantly contributed to Sweet's acquittal. Perhaps the acquittal resulted from the cumulative effect of a multiplicity of factors, including Darrow's repeated salutary racial references (perhaps throughout the trial), the makeup of that particular jury and the dynamics of their group decision-making, the judge's demeanor, and the pre-trial publicity and media coverage. Future research will have to try to parse out each of these variables and weigh its influence on the accuracy and fairness of the factfinding process. My analysis has not aimed to pro-

---

172. *Id.* (citing *Draper v. Airco, Inc.*, 580 F.2d 91, 95 (3d Cir. (1978))). Notwithstanding this court's efforts to compel Professor McClellan and his partner to adopt a colorblind approach to trying cases for black clients, Professor McClellan reports that he and his partner still frequently challenge jurors to be unbiased in judging the claims of their black clients.

173. See, e.g., Douglas L. Colbert, *Challenging the Challenge: Thirteenth Amendment as a Prohibition Against the Racial Use of Preemptory Challenges*, 76 CORNELL L. REV. 1, 110-15 (1990) (discussing empirical findings that all-white juries are not always impartial); Johnson, *supra* note 15, at 1616-51 (discussing data on the influence of racial bias on a determination of guilt).

174. See, e.g., Daniel Coleman, *Study Finds Jurors Often Hear Evidence With a Closed Mind*, N.Y. TIMES, Nov. 29, 1994, at C1. Furthermore, being human, judges are no less subject to unconscious, prejudice-like responses than jurors. Thus, judges might also benefit if rationality-enhancing comments are made during opening statements.

vide a ready answer to these crucial questions, but to provide a theoretical framework for approaching them. Above all, I have attempted to establish that justice often will be better promoted in litigation if we consciously confront stereotypes, than if we take a colorblind, ostrich-head-in-the-sand approach.

#### CONCLUSION

Early in this now elderly century, W.E.B. Du Bois prophesied, in his monumental book, *The Souls of Black Folk*, that "[t]he problem of the twentieth century is the problem of the color-line,—the relation of the darker to the lighter races of men in Asia and Africa, in America and the islands of the sea."<sup>175</sup> Over ninety years later, race continues to be an American obsession and racial discrimination continues to pervade every aspect of the American experience. I have argued that on one level there has been significant progress in race relations since the publication of *Souls of Black Folk*—the percentage of Americans who personally accept and endorse the black stereotype has decreased. On the other hand, as a well-learned set of associations, stereotypes continue to be well-established in the memories of all Americans. Hence, as Professors Lawrence, Johnson, Davis and others have pointed out, we are all prone to stereotype-congruent or prejudice-like responses to blacks (and members of other stereotyped groups), especially in unguarded moments. But research and experience suggest that, in some circumstances, it is possible to resist falling into the discrimination habit. Further progress in eliminating discrimination will require a deeper understanding of the habitual nature of our responses to stereotyped groups and the development of strategies for helping people inhibit their habitual and activate their endorsed responses to these groups.

---

175. W.E.B. DU BOIS, *THE SOULS OF BLACK FOLK* 10 (Bantam Books 1989).