

## **Social Justice in Our Minds, Homes, and Society: The Nature, Causes, and Consequences of Implicit Bias**

**Laurie A. Rudman<sup>1</sup>**

---

*Social injustice is disruptive both personally and collectively, and the ordinariness of implicit prejudice plays a role in this process. In this report, I discuss the correspondence between implicit and explicit biases, and factors that moderate their association. In addition, I differentiate between declarative and procedural definitions of implicit bias, which have implications for their nonconscious nature. To underscore their conceptual distinction, I also present evidence that points to different origins for implicit and explicit biases. Finally, actions for confronting implicit bias are prescribed.*

---

**KEY WORDS:** implicit social cognition; prejudice; stereotypes.

In 1964, when I was 10 years old, my dad drove off with my older sister, Carol. When he returned, my parents announced she would not be coming back. She wanted to marry Lenny—a young African American she had met in college. I couldn't understand what the problem was. The one time he came to our house, I sat on his lap and was thrilled with the way he laughed at my jokes. He was generous and kind and, in fact, years later, he would head the United Way in Chicago. My father had taken Carol to Northeast Minneapolis and told her to make a choice. It was either "those people" or her family. My sister got out of the car.

I knew what prejudice was from television, where I saw Blacks rioting in protest and Martin Luther King, Jr., exhorting all Americans to defeat racism, once and for all. But I had no idea it lived under my own roof until the day I lost my sister. It is even possible that my parents were unaware of their own feelings until they were confronted with my sister's love for Lenny. I never heard

<sup>1</sup>All correspondence should be addressed to Laurie A. Rudman, Department of Psychology, Tillett Hall, Rutgers University, 53 Avenue E, Piscataway, New Jersey, 08854-8040; e-mail: rudman@rci.rutgers.edu.

them speak a prejudiced word, yet her “crime” was dealt with in the harshest terms.

That same year, the Civil Rights Act was passed. It became illegal, as well as immoral, to discriminate against people on the basis of group membership. The result has been an apparent sea change with respect to stereotypes and prejudice, marked by a dramatic decrease in their overt expression (Judd *et al.*, 1995; Schuman *et al.*, 1997), and an increase in normative pressures to be nonprejudiced (Dunton and Fazio, 1997; Plant and Devine, 1998). Indeed, if researchers were to rely solely on self-report measures, they would be tempted to conclude that prejudice has become, if not outdated, at least unfashionable. Clearly, people have had their “consciousness raised” when it comes to reporting how they feel about many social groups (including minorities, women, and the elderly).

However, this does not mean that the problem of bigotry has been solved, for when attitudes are measured using methods that do not rely on respondents’ willingness or ability to report their opinions, the persistence of prejudice and stereotypes is routinely exposed (Dovidio *et al.*, 1997; Fazio *et al.*, 1995). For example, in the most comprehensive report to date, Nosek *et al.* (2002) analyzed data from thousands of Implicit Association Test (IAT) Web site visitors (<http://implicit.harvard.edu/>) and found large effect sizes (i.e., greater than 0.80, using Cohen’s *d*) for implicit biases based on race and age, whereas self-reports yielded much weaker effects (i.e.,  $d_s < 0.52$ ). These findings echo laboratory investigations, which have also uncovered a similar pattern for prejudices based on religion, physical appearance, and socioeconomic class (e.g., Rudman *et al.*, 2002), as well as sexual orientation (e.g., Banse *et al.*, 2001). The hidden nature of prejudice is one of its most pernicious aspects, whether the secret is kept from others or ourselves (Gaertner and Dovidio, 1986). Biases that we do not acknowledge but that persist, unchallenged, in the recesses of our minds, undoubtedly shape our society and may even have played a role in my family’s tragedy.

In this paper I will address three questions, the seeds of which were planted in my childhood when I struggled to understand why my sister was unfairly banished, and that I continue to face as a scholar attempting to shed light on social injustice. First, did my parents know they were prejudiced, or did my sister’s choice of a marital partner uncover largely unconscious feelings? That is, how do explicit and implicit biases relate to each other, and what does it mean to say that we are implicitly biased? Second, how did my parents become biased? That is, what are the underlying causes of implicit bias? Third, what happened in my family in 1964 connects with the larger, contemporaneous struggle against prejudice, but where are we now, and how might organizations and institutions confront implicit bias in order to spur social progress today? Due to space limitations, my coverage will necessarily be selective. Thus, I will focus on prejudice and stereotypes (hereafter, orientations) and on research using priming methods or the IAT (Greenwald *et al.*, 1998), because these are the most prominent techniques

(for more inclusive reviews, see Blair, 2001, 2002; Dovidio *et al.*, 2001; Fazio and Olson, 2003).

### CORRESPONDENCE BETWEEN IMPLICIT AND EXPLICIT MEASURES

Priming and the IAT belong to a class of implicit measures known as response latency methods, in which orientations are inferred from people's ability to complete reaction time tasks. Results have generally revealed pervasive biases that are not mirrored by self-reports (for meta-analyses, see Blair, 2001; Dovidio *et al.*, 2001). For example, Dovidio *et al.* (2001) found modest implicit–explicit (hereafter, I-E) covariation between racial prejudice measures (average  $r = 0.24$ ). Nosek and Banaji (2002) included a wider range of attitudes, but reported similar I-E correspondence (average  $r = 0.29$ ). These findings support conceptualizing implicit and explicit orientations as related but distinct concepts (Blair, 2001; Wilson *et al.*, 2000).

From the above, it appears that self-reports have weak utility when it comes to predicting people's implicit biases. However, in all of the literature reviews, variability in I-E relationships has been the most outstanding feature to emerge. For example, Nosek and Banaji (2002) found virtually no I-E relationship when ageism was assessed ( $r = 0.01$ ), but strong I-E covariation for attitudes toward politicians, whether operationalized as Bush versus Gore ( $r = 0.62$ ), or Democrats versus Republicans ( $r = 0.45$ ). Similarly, Blair (2001) reported I-E prejudice linkages ranging from  $r = -0.28$  (Whites vs. Blacks; Fazio *et al.*, 1995) to  $r = 0.58$  (Korean vs. Japanese; Greenwald *et al.*, 1998). Using the identical measures, Dovidio *et al.* (1997) found a strong relationship ( $r = 0.60$ ) between race priming and the Modern Racism Scale (MRS; McConahay, 1986) in Study 2, but no relationship in Study 3 ( $r = 0.01$ ). Similarly, the links between race priming and the MRS that have been found in the past ( $r_s > 0.39$ ; Wittenbrink *et al.*, 1997, 2001) have failed to replicate (Boniecki and Jacks, 2002; Rudman *et al.*, 2001). These agonizingly disparate findings have prompted authors to echo Blair's (2001) conclusion that "... simple answers to the question of what implicit measures reveal about an individual's attitude or belief (as understood through explicit measures) are not forthcoming" (p. 362).

### Moderators of I-E Correspondence

However, as noted by Fazio and Olson (2003), the question is not *whether* there is I-E convergence, but what factors moderate their relationship? Using measures of racial prejudice, stronger I-E covariation has been found among respondents low on motives to be nonprejudiced, compared with those who are

high (Dunton and Fazio, 1997; Fazio *et al.*, 1995). Similar results have emerged when examining heterosexism (Banse *et al.*, 2001; Dasgupta and Rivera, 2003, Exp. 2). Thus, motives to control prejudice can sometimes moderate I-E prejudice correspondence (see also Dasgupta, 2004).

More generally, Nosek and Banaji (2002) argued that I-E linkages depend on self-presentation motives. For example, racial attitudes are likely to be self-censored more than attitudes toward academic subjects. Using 15 attitude objects, they asked an independent sample to rate the likelihood that each would evoke self-presentation concerns; these estimates predicted I-E correspondence in a secondary analysis ( $\beta = 0.84, p < 0.01$ ). Additionally, they found support for their reasoning that elaborated attitude objects would elicit greater I-E correspondence,  $\beta = 0.61, p < 0.05$ . For example, people are likely to think more often and deeply about presidential candidates than apples and candy bars. This possibility helps to explain why attitude objects that are low in self-presentation concerns (e.g., apples and candy bars) can result in low I-E correspondence (Karpinski and Hilton, 2001). Although Nosek and Banaji's results require replication, self-presentation motives and cognitive elaboration appear to be promising moderator variable candidates.

Additional moderator variables have been suggested, including attitude accessibility and importance (Nosek and Banaji, 2002). Further, Greenwald and Nosek (2001) noted that restricted range on either measure inhibits I-E convergence. Given the tendency to observe floor effects on explicit prejudice measures, this problem is likely to be persistent. Finally, reliability issues are important for I-E covariation. Although the IAT has shown excellent internal consistency and adequate test-retest stability, priming measures tend to suffer from substantial noise (cf. Fazio and Olson, 2003). For example, one study found a stability coefficient (over a 10-week period) of 0.47 for the IAT, but 0.08 for sequential priming (Rudman *et al.*, 2001, Exp. 2). Thus, psychometric properties, as well as theoretical variables, must be considered when I-E linkages are examined (Cunningham *et al.*, 2001; Greenwald *et al.*, 2003).

### Validation of Implicit Measures

The absence of consistently strong I-E convergence underscores the discriminant validity of response latency techniques, but it also requires researchers to validate them through other means. To date, known groups validity has been routinely demonstrated, and support for their predictive utility has steadily accrued (see Fazio and Olson, 2003; Greenwald and Nosek, 2001, for reviews). For example, prosocial behavior toward Blacks is negatively predicted by implicit prejudice, whether prejudice is assessed by the IAT (Ashburn-Nardo *et al.*, 2003; McConnell and Leibold, 2001; Rudman and Lee, 2002) or priming (Dovidio *et al.*, 1997, 2002; Fazio *et al.*, 1995). In addition, discriminating against female job applicants is predicted by implicit stereotypes, whether stereotypes are assessed by

priming (Rudman and Borgida, 1995) or the IAT (Rudman and Glick, 2001). Taken together, the evidence supports response latency measures as indicators of individual differences in implicit orientations, and demonstrates that these biases have consequences for equal opportunity (see also Dasgupta, 2004).

## THE NATURE OF IMPLICIT ORIENTATIONS

Having established that implicit biases (1) detect aspects of social reality that are hidden by self-reports and (2) may guide important judgments and behaviors, in this section I will address two questions—What are they? What are their causes? Because implicit social cognition is a young field, my answers are necessarily preliminary and are offered in the spirit of theoretical advance.

### What are Implicit Biases?

Prejudice can be defined as an association in long-term memory between social objects and valence (positive or negative; Fazio, 1990); similarly, stereotypes are thought to be well-learned associations between social objects and attributes. In each case, the orientation is routinized to the point where it is automatically accessed in the presence of objects. Because people can control their explicit (but not implicit) responses, response latency measures yield estimates of bias that are comparatively “upstream” (i.e., closer to the source). That is, even when people are truthful, self-reports can *only* reflect what they believe about their orientations, whereas implicit measures bypass this limitation. Following established parameters for automaticity (Bargh, 1989), implicit biases are thought to be automatic not only in the sense that they are fast-acting, but also because they can operate without (1) intention (i.e., are involuntary and uncontrollable), and (2) conscious awareness. For this reason, implicit biases have also been described as automatic or nonconscious (e.g., Blair, 2001).

The frequent observation of I-E dissociation suggests that implicit orientations are indeed unintentional, because people generally strive to be consistent in their responses (Festinger, 1957). In addition, people cannot fake their implicit attitudes (e.g., Banse *et al.*, 2001; Egloff and Schmukle, 2002), in support of their uncontrollability. But whether or not they are nonconscious is a point of debate (Fazio and Olson, 2003). Unfortunately, we cannot be sure that implicit biases are nonconscious, due to the difficulty of assessing people’s awareness independent of asking them (see also Greenwald *et al.*, 2002a). We do know that while many Whites can subjectively tell that it is easier for them to associate Blacks with negative, as opposed to positive, evaluation on the race IAT, they resist attributing this proclivity to prejudice (Monteith *et al.*, 2001). Of course, this is a far cry from demonstrating that implicit biases are inaccessible. More promisingly, theories

that posit nonconscious processes have been tested using implicit and explicit measures. When implicit (but not explicit) responses support the theory, it may well be because implicit biases are nonconscious (Dovidio *et al.*, 1997; Jost *et al.*, 2002; Rudman *et al.*, 2002). Indeed, when people are truly unable to access their automatic associations, only implicit measures can detect them. Nonetheless, the use of the term “nonconscious” to describe implicit biases has become a sticking point. As discussed below, the point of debate centers on whether the definition is declarative (describing mental contents) or procedural (describing how implicit biases operate).

In the declarative camp, implicit and explicit attitudes (including prejudice) are both defined as well-learned associations between objects and evaluations (Fazio, 1990; Fazio *et al.*, 1995). From this perspective, it seems contradictory to claim that routinized associations are inaccessible to introspection; after all, people had to learn them at some point. However, it is possible that people might change a previously learned attitude, but that both coexist in long-term memory, rather than the new evaluation replacing the old. The result would be a “dual-attitude,” with the old attitude preserved in nonconscious form (Wilson *et al.*, 2000). Another possibility is that people might not know they have learned an implicit orientation (i.e., the learning might have occurred incidentally). For example, stereotypic media portrayals might classically condition people’s implicit attitudes without their being aware of this effect (cf. Olson and Fazio, 2002). Nonetheless, because they share the same declarative definition, implicit and explicit biases have been argued to be more distinguishable methodologically than conceptually, and aspersions have been cast on the nonconscious nature of implicit associations (Fazio and Olson, 2003).

In the procedural camp, Greenwald and Banaji (1995) theorized that implicit attitudes are “introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects” (p. 8). Similarly, Wilson *et al.* (2000) defined implicit attitudes as “evaluations that (a) have an unknown origin (people are unaware of the basis of their evaluation); (b) are activated automatically; and (c) influence implicit responses, namely, uncontrollable responses and ones that people do not view as an expression of their attitude and thus do not attempt to control” (p. 104). Note that these definitions do not state that the evaluation is nonconscious but, instead, that its source is not consciously considered as a causal influence. For this reason, judgments and behavior may be influenced by implicit orientations without intention or awareness. That is, the *application* of implicit biases may be nonconscious.

An example serves to illustrate the point. Men primed with commercials portraying women as sex objects subsequently (1) showed more automatic activation of a negative female subtype, and (2) behaved toward a female job applicant in a more sexist manner, compared with controls (Rudman and Borgida, 1995). During debriefings, men reported being unaware of the influence of the primes (resulting

in an inability to check their behavior). However, men are likely aware that they are capable of thinking about women as sex objects. What is *implicit* is the effect of the priming manipulation on their subsequent thoughts and actions (see also Devine, 1989; Rudman and Lee, 2002).

In sum, if we declaratively define implicit and explicit orientations as “well-learned associations” but also specifically define the former in procedural terms, we need not require that people be unaware of implicit biases when we distinguish automatic from controlled responses. Although it is entirely feasible that people are, at least at times, unable to access the content of their minds (Wilson *et al.*, 2000), it may be more prudent to claim that people are unaware of (1) the influence that the source (e.g., traces of past experience) has on their implicit biases, and (2) the influence of their implicit biases on their judgments and behavior. That is, the source and the impact of the evaluation can be nonconscious without the evaluation itself being inaccessible. In the next section, I will outline potential sources of implicit bias. In the last section, I will address prescriptions for the unwitting application of implicit biases.

### What Causes Implicit Biases?

Implicit and explicit orientations are likely to be distinguishable, in part, by their disparate causes (Rudman, 2003). Compared with explicit opinions, implicit biases may be more influenced by early experiences, affective experiences, cultural biases, and cognitive balance principles. Below, I will briefly present the supporting evidence, and speculate that a fifth source (the self) may also be key.

First, it has been theorized that implicit orientations stem from past (e.g., developmental) experiences, whereas explicit orientations may reflect more recent events (Greenwald and Banaji, 1995; Wilson *et al.*, 2000). Consistent with this view, Rudman and Goodwin (2003) found that people’s earliest experiences with females (maternal caregivers) influenced automatic gender attitudes. First, people raised primarily by their mothers automatically preferred women to men. Second, implicit parent and gender attitudes were linked, such that people favored women over men if they also preferred their mothers to their fathers. By contrast, explicit parent and gender attitudes did not covary. Thus, positive experiences with a primary female object biased attitudes toward women in general, but only at the automatic level. In addition, Rudman and Heppen (2003) found that women showed strong possession of implicit romantic fantasies (i.e., association of male partners with Prince Charming and White Knight) that was not mirrored by their self-reports. Further, implicit (but not explicit) fantasies negatively predicted women’s interest in personal power (e.g., choosing a leadership role). Because women are socialized early (and often) to view men as their heroes (e.g., through romantic fairy tales), these results indirectly support the hypothesis that developmental events can influence automatic mental habits.

Second, implicit biases may be more sensitive to affective experiences, compared with self-reports. For example, implicit (but not explicit) anti-Black prejudice covaried with amygdala activation in Whites exposed to photos of Blacks (Phelps *et al.*, 2000). Because the amygdala is associated with emotional learning, including fear conditioning, the inference is that implicit biases are linked to perceptions of anxiety or threat. In addition, Rudman *et al.* (2001) found that Whites who volunteered for diversity education showed reduced anti-Black implicit and explicit prejudice and stereotypes at the end of a 12-week course. However, I-E change scores were weakly associated. Further analyses revealed that reductions in implicit orientations were predicted by emotion-based factors, including reduced fear of Blacks, and liking for the Black professor who taught the course. By contrast, reductions in explicit orientations covaried with students' awareness of (and motives to overcome) their own prejudice, brought about as a result of the course. These findings suggest that changes at the automatic level may depend on emotional reconditioning. Finally, Rudman and Goodwin (2003) found that, compared with controls, participants exposed to information about violent White male criminals showed an increase in their implicit (1) male-threat associations (i.e., they more readily associated men with violence and threat) and (2) profemale evaluation. By contrast, explicit gender attitudes and beliefs were insensitive to the threat manipulation (see also Florack *et al.*, 2001; Towles-Schwen and Fazio, 2003, for related research).

Third, implicit orientations may be more influenced by cultural biases, compared with explicit attitudes and beliefs (Devine, 1989). For example, high status groups typically show stronger automatic ingroup bias, compared with low status groups, whether status is based on ethnicity, age, religion, physical attractiveness, SES, or college affiliation (Jost *et al.*, 2002; Nosek *et al.*, 2002; Rudman *et al.*, 2002). Because high status groups are culturally favored, these findings suggest that societal appraisals permeate the implicit cognitions of all group members (Jost and Banaji, 1994). In support of this hypothesis, dominant and subordinate groups alike showed covariation between perceived status differences and implicit (but not explicit) ingroup bias (Rudman *et al.*, 2002). That is, people preferred the dominant group to the extent that it was culturally favored, but only at the automatic level. Interestingly, this pattern breaks down when it comes to gender attitudes, because women show stronger implicit ingroup bias than do men (due, in part, to mitigating sources; Rudman and Goodwin, 2003). However, men and women alike automatically preferred male to female authority figures (e.g., doctors), but particularly if they associated men more than women with status (Rudman and Kilianski, 2000). Thus, status may interact with social role when automatic gender evaluations are at stake (Eagly and Karau, 2002).

Finally, implicit orientations are influenced by cognitive balance principles (Greenwald *et al.*, 2002a). For example, people who liked themselves *and* identified with their ingroup also showed robust ingroup bias, whereas self-report measures showed no hint of cognitive consistency. The dependence of ingroup evaluation

on the interaction of self-appraisal and group identity is the hallmark of the unified theory of implicit social cognition, which has provided some of the most compelling evidence that implicit and explicit orientations are derived from different sources. The pattern can be characterized as “If I am good and I am X, then X is also good” (where X represents group identification). For example, Whites with high self-esteem who strongly identified as White also preferred Whites to Blacks on the race IAT.

Similar findings emerged when evaluation and self-esteem were replaced by stereotypes and self-concept, respectively, which supports the generality of the unified model. Interestingly, the resulting pattern can be counterstereotypical (e.g., “If I am warm and I am male, then men are warmer than women”), suggesting that the self may be implicated in automatic associations more than cultural milieu. In fact, as I have speculated elsewhere (Rudman, 2003), it may be difficult to possess implicit associations that are dissociated from the self, whereas controlled evaluations may allow for more objective responses. If the self proves to be a central cause of implicit orientations, it is likely because we do not view ourselves impartially, and this partisanship then shapes appraisals of other objects that are (or are not) connected to the self. Although Greenwald *et al.* (2002a) were agnostic about motives underlying the unified model, subsequent research has underscored the dramatic ease with which people align themselves with novel groups (Greenwald *et al.*, 2002b), and assimilate novel stimuli into existing self-relevant groups (Ashburn-Nardo *et al.*, 2001), resulting in implicit attitudes that might best be accounted for by the self’s central role in automatic responses. Because, on average, people possess robust implicit self-esteem (Nosek *et al.*, 2002) and strongly identify with many social groups (Greenwald *et al.*, 2002b), self-partisanship should not be overlooked as a potentially potent cause of implicit orientations.

In sum, a consideration of the sources of implicit biases supports their theoretical distinction from explicit biases, and also offers another reason why I-E correspondence may be weak. When implicit and explicit biases stem from disparate causes, we should not expect their underlying evaluation to be the same. That is, the “well-learned associations” that define automatic and controlled orientations might even be opposite in valence if they spring from different wells (Wilson *et al.*, 2000). Because past research has largely focused on process differences between automatic and controlled responses, the tendency has been to blame their (often-observed) dissociation on the untrustworthiness of self-reports. In contrast, examining their bases provides a rationale for deeming both types of responses as potentially valid, albeit limited in their ability to encompass the range of human responses to social objects.

## WHAT SHOULD WE DO ABOUT IMPLICIT BIASES?

Response latency measures have exposed the unhappy ordinariness of implicit biases, and the extent to which they can (unintentionally) guide our thoughts and

actions. Biases based on race, nationality, religion, class, age, sex, and sexual orientation (to cite a few examples) have come to light when people are unable to control their responses. Moreover, people may be involuntarily influenced by these biases when they make evaluative (e.g., hiring) decisions, or when they interact with others in ways that can lead to discomfort or misunderstanding. What should we do about this state of affairs?

As a first step, we need to inform people about this picture of social reality. Because awareness of a problem is the first step to eliminating it (Wilson and Brekke, 1994), education is key. There has been, and will be, backlash for these efforts (see Banaji, 2001), but they are critical to our ability to provide descriptions of social cognition that are faithful to human complexity, as well as to endeavors to combat automatic prejudice.

Second, implicit biases should make us double our efforts to retain and expand policies that serve to override them, including affirmative action. This is particularly important given the current administration, which, in the field of education, is trying to substitute policies that are likely to reduce opportunities for minorities while simultaneously lowering academic standards. Although the Supreme Court recently upheld the University of Michigan's Law School efforts to diversify its student body, the university's policy was struck down, a decision that can only fuel the growing backlash against affirmative action (Cokerinos, 2003). Moreover, Justice O'Connor stated in her decision that affirmative action has a definite shelf life (25 years). In light of the prevalence of implicit biases and their consequences for equal opportunity, this assessment is overly optimistic.

Third, we should keep in mind the connection between self and cultural milieu (Banaji, 2001). To the extent that societal evaluations color implicit biases, a more inclusive society should reduce them. As an explanation for why he automatically favored Whites, a Black student quickly observed, "We've never had a Black president." Similarly, the absence of visible female authority figures can help to preserve gender hegemony (Rudman and Kilianski, 2000). Indeed, recent evidence suggests that women who attend women's colleges reduced their implicit gender stereotypes (linking men to leadership and women to supportive roles) in the course of a year precisely because they were exposed to more female faculty and deans, compared with their co-ed counterparts (Dasgupta and Asgari, 2003). Thus, organizations should strive to increase the representation of minorities and women in positions of power and prestige.

Fourth, it is an axiom that proximity leads to attraction; indeed, providing people with opportunities to become emotionally comfortable with outgroup members can reduce implicit biases (Rudman *et al.*, 2001). However, although the idea that intergroup contact promotes egalitarianism is a venerable one, contact can have a *negative* effect on implicit orientations (Livingston, 2002) unless it promotes mutual trust. Organizations and institutions cannot force people to interact, but more could be done to curb their tendencies to segregate themselves (Anderson

and Massey, 2001; Moody, 2001) and to foster the type of interactions that are likely to decrease perceptions of intergroup threat and, thereby, implicit biases (Phelps *et al.*, 2000; Rudman *et al.*, 2001).

Finally, we should be realistic about how much we can change people's automatic responses. We may never eliminate them, particularly to the extent they reflect self-partisanship. From the cradle, people are raised by nurturers who inject their values with an irresistible dose of affection, which inevitably leads them to identify ingroup members (family and friends) before recognizing outsiders (Allport, 1954). As a result, it is perfectly natural to favor those who affirm our values and to devalue others who, by their very existence, seemingly threaten our way of life. That is, a preference for one's ingroup may be as ordinary as a preference for one's own children; although only the former is unethical (and, in some circumstances, illegal), both can be viewed as a symptom of the same causal root. What we *can* do is become aware of the power of implicit partisanship. Thus, the first and final recommendation is education, or what John Jost has dubbed "(un)consciousness raising" (Banaji, 2001, p. 136). To succeed, we may have to break down some of our defenses, for we are well practiced at stiff-arming negative information about ourselves (e.g., Greenwald, 1980; Monteith *et al.*, 2001). Although this habit may protect our self-esteem, it can have negative consequences for society's well-being, which affects all of us, both collectively and personally.

### POSTSCRIPT

I began this paper with my own story of incalculable loss because it illustrates the pain that prejudice causes not just on a societal level, but also on a deeply personal one. Social injustice deprived me of a loving sister whose memory shines bright to this day, and of an extended family that I would have gladly welcomed. This was compounded by the humiliation I felt upon learning that my parents were racists. And perhaps that is a hopeful message. For in one generation, the stigma changed from being black to being bigoted; even in my 10-year-old mind, it was the latter that was shameful. And while the fallout from racism continues to plague our society, the fact that Americans often disavow prejudice is a promising turn of events when compared with the overt biases exhibited in the past. However, for a deep and lasting equality to evolve, implicit biases must be acknowledged and challenged; to do otherwise is to allow them to haunt our minds, our homes, and our society into the next millennium.

### ACKNOWLEDGMENTS

This report was supported by Grant BCS-0109997 from National Science Foundation. I thank Dolly Chugh and Nilanjana Dasgupta for their comments on an earlier draft.

## REFERENCES

- Allport, G. W. (1954). *The Nature of Prejudice*, Addison-Wesley, New York.
- Anderson, E., and Massey, D. S. (2001). *Problem of the Century: Racial Stratification in the United States*, Russell Sage Foundation, New York.
- Ashburn-Nardo, L., Knowles, M. L., and Monteith, M. J. (2003). Black Americans' implicit racial associations and their implications for intergroup judgment. *Soc. Cogn.* 21: 61–87.
- Ashburn-Nardo, L., Voils, C. I., and Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *J. Pers. Soc. Psychol.* 81: 789–799.
- Banaji, M. R. (2001). Implicit attitudes can be measured. In Roediger, H. L., Nairne, J. S., Neath, I., and Surprenant, A. M. (eds.), *The Nature of Remembering: Essays in Honor of Robert G. Crowder*, American Psychological Association, Washington, DC, pp. 117–150.
- Banse, R., Seise, J., and Zerbes, N. (2001). Implicit attitudes toward homosexuality: Reliability, validity, and controllability of the IAT. *Z. Exp. Psychol.* 48: 1145–1160.
- Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In Uleman, J. S., and Bargh, J. A. (eds.), *Unintended Thought*, Guilford, New York, pp. 3–51.
- Blair, I. V. (2001). Implicit stereotypes and prejudice. In Moskowitz, G. B. (ed.), *Cognitive Social Psychology: The Princeton Symposium on the Legacy and Future of Social Cognition*, Erlbaum, Mahwah, NJ, pp. 359–374.
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Pers. Soc. Psychol. Rev.* 6: 242–261.
- Bonietti, K. A., and Jacks, J. Z. (2002). The elusive relationship between measures of implicit and explicit prejudice. *Representative Res. Soc. Psychol.* 26: 1–14.
- Cokerinos, L. (2003). *The Assault on Diversity: An Organized Challenge to Racial and Gender Justice*, Rowman and Littlefield, Lanham, MD.
- Cunningham, W. A., Preacher, K. J., and Banaji, M. R. (2001). Implicit attitude measures: Consistency, reliability, and convergent validity. *Psychol. Sci.* 121: 163–170.
- Dasgupta, N. (2004). Implicit ingroup favoritism, outgroup favoritism, and their behavioral manifestations. *Soc. Justice Res.* 17(2): 143–168.
- Dasgupta, N., and Asgari, S. (2003). *Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping and self-stereotyping*. Manuscript submitted for review.
- Dasgupta, N., and Asgari, S. (in press). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *J. Exp. Soc. Psychol.*
- Dasgupta, N., and Rivera, L. (2003). *The road sometimes taken: How negative affect, perceiver sex, and target sex influence the path from implicit sexual prejudice to behavior*. Manuscript in preparation.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *J. Pers. Soc. Psychol.* 56: 5–18.
- Dovidio, J. F., Kawakami, K., and Beach, K. R. (2001). Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias. In Brown, R., and Gaertner, S. L. (eds.), *Blackwell Handbook of Social Psychology, Vol. 4: Intergroup Relations*, Blackwell, Oxford, UK, pp. 175–197.
- Dovidio, J. F., Kawakami, K., and Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interactions. *J. Pers. Soc. Psychol.* 82: 62–68.
- Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., and Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *J. Exp. Soc. Psychol.* 33: 510–540.
- Dunton, B. C., and Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Pers. Soc. Psychol. Bull.* 23: 316–326.
- Eagly, A. H., and Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychol. Rev.* 109: 573–598.
- Egloff, B., and Schmukle, S. C. (2002). Predictive validity of the Implicit Association Test for assessing anxiety. *J. Pers. Soc. Psychol.* 83: 1441–1455.
- Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In Zanna, M. P. (ed.), *Advances in Experimental Social Psychology*, Vol. 23, Academic Press, New York, pp. 75–109.

- Fazio, R. H., Jackson, J. R., Dunton, B. C., and Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes. A bona fide pipeline? *J. Pers. Soc. Psychol.* 69: 1013–1027.
- Fazio, R. H., and Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Ann. Rev. Psychol.* 54: 297–327.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Row, Peterson and Company, Evanston, IL.
- Florack, A., Scarabis, M., and Bless, H. (2001). When do associations matter? The use of associations toward ethnic groups in person judgments. *J. Exp. Soc. Psychol.* 37: 518–524.
- Gaertner, S. L., and Dovidio, J. F. (1986). The aversive form of racism. In Dovidio, J. F., and Gaertner, S. L. (eds.), *Prejudice, Discrimination, and Racism*, Academic Press, Orlando, FL, pp. 61–90.
- Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *Am. Psychol.* 35: 603–618.
- Greenwald, A. G., and Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychol. Rev.* 102: 4–27.
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., and Mellott, D. S. (2002a). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychol. Rev.* 109: 3–25.
- Greenwald, A. G., McGhee, D., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *J. Pers. Soc. Psychol.* 74: 1464–1480.
- Greenwald, A. G., and Nosek, B. A. (2001). The health of the Implicit Association Test at age 3. *Z. Exp. Psychol.* 48: 85–93.
- Greenwald, A. G., Nosek, B. A., and Banaji, M. R. (2003). Understanding and using the Implicit Association Test: An improved scoring algorithm. *J. Pers. Soc. Psychol.* 85: 197–216.
- Greenwald, A. G., Pickrell, J. E., and Farnham, S. D. (2002b). Implicit partisanship: Taking sides for no reason. *J. Pers. Soc. Psychol.* 83: 367–379.
- Jost, J. T., and Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false-consciousness. *Br. J. Soc. Psychol.* 33: 1–27.
- Jost, J. T., Pelham, B. W., and Carvallo, M. R. (2002). Non-conscious forms of system justification: Implicit and behavioral preferences for higher status groups. *J. Exp. Soc. Psychol.* 38: 586–602.
- Judd, C. M., Park, B., Ryan, C. S., Brauer, M., and Kraus, S. (1995). Stereotypes and ethnocentrism: Interethnic perceptions of African American and White American college samples. *J. Pers. Soc. Psychol.* 69: 460–481.
- Karpinski, A., and Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *J. Pers. Soc. Psychol.* 81: 774–788.
- Livingston, R. W. (2002). The role of perceived negativity in the moderation of African Americans' implicit and explicit racial attitudes. *J. Exp. Soc. Psychol.* 38: 405–413.
- McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In Dovidio, J. F., and Gaertner, S. L. (eds.), *Prejudice, Discrimination, and Racism*, Academic Press, Orlando, FL, pp. 91–126.
- McConnell, A. R., and Leibold, J. M. (2001). Relations among the Implicit Association Test, explicit attitudes, and discriminatory behavior. *J. Exp. Soc. Psychol.* 37: 435–442.
- Monteith, M. J., Voils, C. I., and Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit biases. *Soc. Cogn.* 19: 395–417.
- Moody, J. (2001). Race, school integration, and friendship segregation in America. *Am. J. Soc.* 107: 679–716.
- Nosek, B. A., and Banaji, M. R. (2002). (At least) two factors moderate the relationship between implicit and explicit attitudes. In Ohme, R. K., and Jarymowicz, M. (eds.), *Natura Automatyzmow*, WIP PAN and SWPS, Warszawa, pp. 49–56.
- Nosek, B. A., Banaji, M. R., and Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dyn. Theory Res. Pract.* 6: 101–115.
- Olson, M. A., and Fazio, R. H. (2002). Implicit acquisition and manifestation of classically conditioned attitudes. *Soc. Cogn.* 20: 89–104.
- Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., and Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *J. Cogn. Neurosci.* 12: 729–738.

- Plant, E. A., and Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *J. Pers. Soc. Psychol.* 75: 811–832.
- Rudman, L. A. (2004). Sources of implicit attitudes. *Curr. Dir. Psychol. Sci.* 13(2): 80–83.
- Rudman, L. A., Ashmore, R. D., and Gary, M. L. (2001). “Unlearning” automatic biases: The malleability of implicit stereotypes and prejudice. *J. Pers. Soc. Psychol.* 81: 856–868.
- Rudman, L. A., and Borgida, E. (1995). The afterglow of construct accessibility: The behavioral consequences of priming men to view women as sexual objects. *J. Exp. Soc. Psychol.* 31: 493–517.
- Rudman, L. A., Feinberg, J. M., and Fairchild, K. (2002). Minority members’ implicit attitudes: Ingroup bias as a function of group status. *Soc. Cogn.* 20: 294–320.
- Rudman, L. A., and Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *J. Soc. Issues* 57: 743–762.
- Rudman, L. A., and Goodwin, S. A. (2003). [Priming male violence increases implicit pro-female bias]. Unpublished raw data.
- Rudman, L. A., and Goodwin, S. A. (in press). *Gender differences in automatic in-group bias: Why do women like women more than men like men?* *J. Pers. Soc. Psychol.*
- Rudman, L. A., and Heppen, J. (2003). Implicit romantic fantasies and women’s interest in personal power: A glass slipper effect? *Pers. Soc. Psychol. Bull.* 29: 1357–1370.
- Rudman, L. A., and Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Pers. Soc. Psychol. Bull.* 26: 1315–1328.
- Rudman, L. A., and Lee, M. R. (2002). Implicit and explicit consequences of exposure to violent and misogynous rap music. *Group Process. Intergroup Relat.* 5: 133–150.
- Schuman, H., Steeh, C., Bobo, L., and Krysan, M. (1997). *Racial Attitudes in America: Trends and Interpretations*. Harvard University Press, Cambridge, MA.
- Towles-Schwen, T., and Fazio, R. H. (2003). Choosing social situations: The relation between automatically activated racial attitudes and anticipated comfort interacting with African Americans. *Pers. Soc. Psychol. Bull.* 29: 170–182.
- Wilson, T. D., and Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychol. Bull.* 116: 117–142.
- Wilson, T. D., Lindsey, S., and Schooler, T. Y. (2000). A model of dual attitudes. *Psychol. Rev.* 107: 101–126.
- Wittenbrink, B., Judd, C. M., and Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *J. Pers. Soc. Psychol.* 72: 262–274.
- Wittenbrink, B., Judd, C. M., and Park, B. (2001). Evaluative versus conceptual judgments in automatic stereotyping and prejudice. *J. Exp. Soc. Psychol.* 37: 244–252.

Copyright of Social Justice Research is the property of Kluwer Academic Publishing and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.