CHAPTER 2

Aversive Racism—How Unconscious Bias Influences Behavior: Implications for Legal, Employment, and Health Care Contexts

John F. Dovidio, Samuel L. Gaertner, Louis A. Penner, Adam R. Pearson, and Wynne E. Norton

Over 100 years ago, in his classic book *The Souls of Black Folk,* W.E.B. DuBois (1903/1986) observed: "The problem of the twentieth century is the problem of the color line" (p. 372). On the one hand, consistent declines in the overt expression of racial prejudice among whites and increasing endorsement of egalitarianism as a central cultural value (Bobo, 2001) seem to challenge the validity of DuBois's proposition. Indeed, the nomination of Barack Obama for president of the United States by a major political party and his eventual election victory are landmark events that shatter a major historical racial barrier. On the other hand, significant and pervasive economic and social disparities between blacks and whites in the United States make DuBois's observation prophetic.

In this chapter, we explore the current status and psychological nature of racial bias in the United States. We propose that current racial attitudes of whites toward blacks in the United States are fundamentally ambivalent, characterized by a widespread contemporary form of bias associated with aversive racism, and we consider its influence on the behavior of whites toward blacks. The central question we address is, how can well-intentioned people inadvertently and subtly discriminate to contribute to racial disparities? We hypothesize that the ambivalence reflected in aversive racism can offer broader insights for understanding racial disparities, bias in interracial interactions, and ultimately race relations.

In this chapter, we mainly focus on white-black relations in the United States. We recognize that other forms of intergroup relations, both within the United States and internationally, merit theoretical and empirical attention, particularly when considering the changing demographics in the United States and trends in international immigration. In addition,

within white-black relations, our emphasis is on the impact of whites' attitudes on race relations. We acknowledge that the attitudes of blacks are also important, but because whites have the major share of resources and social power in U.S. society, their attitudes are particularly influential for causing, and potentially ameliorating, racial disparities. In this chapter, we first discuss the existence of a contemporary form of racial attitudes and aversive racism among whites. We then explain how aversive racism is manifested in subtle forms of discrimination that can have significant consequences for outcomes for blacks, as well as for the dynamics of interracial interactions. After that, we explore the implications of aversive racism for understanding racial disparities in legal, employment, and health care contexts. We conclude by reviewing ways of combating the subtle bias associated with aversive racism.

THE NATURE OF AVERSIVE RACISM

Research from the 1920s through the 1950s typically portrayed prejudice as psychopathology (Dovidio, 2001). However, stimulated by developments in the area of social cognition, by the mid 1960s and early 1970s, much more emphasis was devoted to recognizing how *normal* cognitive (e.g., social categorization), motivational (e.g., needs for status), and sociocultural (e.g., social transmission of stereotypes) processes contribute to the development of whites' biases toward blacks (see Dovidio, 2001; Dovidio & Gaertner, 2004). For instance, people automatically distinguish others on the basis of race, and this social categorization spontaneously activates more positive feelings and beliefs about in-group members ("we's") than out-group members ("they's") (see Gaertner & Dovidio, 2000, for a review). In addition, whites automatically activate stereotypes of whites as intelligent, successful, and educated, and of blacks as aggressive, impulsive, and lazy (Blair, 2001). Intergroup processes, such as system-justifying ideologies as well as perceived competition over material resources, can also form a basis for negative racial attitudes (see Sidanius & Pratto, 1999).

Kovel (1970) distinguishes between dominative racism, which is the old-fashioned, blatant form, and aversive racism. Aversive racism reflects a fundamental conflict between whites' denial of personal prejudice and the negative feelings toward and beliefs about blacks, which may be unconscious, that result in the normal psychological processes that promote racial bias (Gaertner & Dovidio, 1986). Because of current cultural values in the United States, most whites have strong convictions concerning fairness, justice, and racial equality. Overt expressions of prejudice in the United States have declined significantly over the past 40 years. As Bobo (2001) concludes in his review of trends in racial attitudes, "the single clearest trend in studies of racial attitudes has involved a steady and sweeping movement toward general endorsement of the principles of racial equality and integration" (p. 269). However, because of the range

of normal psychological processes that promote intergroup biases, most whites also develop unconscious negative beliefs and attitudes about blacks, as well as negative feelings. However, the negative feelings that aversive racists have toward blacks do not reflect open hostility or hatred. Instead, the reactions of aversive racists may involve discomfort, anxiety, or fear. That is, they find blacks "aversive," while at the same time they find any suggestion that they might be prejudiced "aversive" as well.

Aversive racists are presumed to have egalitarian conscious, or explicit, attitudes but negative unconscious, or implicit, racial attitudes (Dovidio & Gaertner, 2004). Explicit attitudes operate in a conscious mode and are exemplified by traditional, self-report measures. Implicit attitudes, in contrast, are evaluations and beliefs that are automatically activated by the mere presence (actual or symbolic) of the attitude object. They commonly function in an unconscious fashion. Implicit attitudes and stereotypes are typically assessed using response latency procedures, memory tasks, physiological measures (e.g., galvanic skin response), and indirect self-report measures (e.g., involving attributional biases). The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), for example, taps the different stereotypic or evaluative (e.g., good-bad) associations that people have with racial groups (but of which they may be unaware) on the basis of the general finding that people make decisions about pairs of stimuli that are similar in valence faster than they make decisions about pairs that differ in valence.

Techniques for assessing implicit attitudes are particularly useful for distinguishing between aversive racists, who endorse egalitarian views and nonprejudiced ideologies but harbor implicit racial biases, and nonprejudiced people, who also endorse egalitarian values but do not have significant implicit prejudice or stereotypes. Consistent with the aversive racism framework, for example, the majority of whites in the United States appear nonprejudiced on self-report (explicit) measures of prejudice, whereas a very large proportion of whites also demonstrate implicit racial biases. Overall, whites' generally negative implicit attitudes are largely dissociated from their typically more positive overt expressions of their attitudes toward blacks (Dovidio, Kawakami, & Beach, 2001). Thus, a substantial portion of whites are characterized by aversive racism.

CONSEQUENCES OF AVERSIVE RACISM

In contrast to the traditional form of racism, which is blatant and expressed openly and directly, aversive racism operates in subtle and indirect ways. Specifically, whereas old-fashioned racists exhibit a direct and overt pattern of discrimination, the actions of aversive racists may appear more variable and inconsistent. Sometimes they discriminate (manifesting their negative feelings), and sometimes they do not (reflecting their egalitarian beliefs). Nevertheless, their discriminatory behavior is predictable.

Because aversive racists consciously recognize and endorse egalitarian values and because they truly aspire to be nonprejudiced, they will *not* act inappropriately in situations with strong social norms when discrimination would be obvious to others and to themselves. Specifically, when they are presented with a situation in which the normatively appropriate response is clear, in which right and wrong are clearly defined, aversive racists will not discriminate against blacks. In these contexts, aversive racists will be especially motivated to avoid feelings, beliefs, and behaviors that could be associated with racist intent. Wrongdoing of this type would directly threaten their nonprejudiced self-image. However, aversive racists still possess unconscious negative feelings and beliefs, which will eventually be expressed but in subtle, indirect, and rationalizable ways. For instance, discrimination will occur in situations in which normative structure is weak, when the guidelines for appropriate behavior are vague, or when the basis for social judgment is ambiguous. In addition, discrimination will occur when an aversive racist can justify or rationalize a negative response on the basis of some factor other than race. Under these circumstances, aversive racists may engage in behaviors that ultimately harm blacks but in ways that allow whites to maintain their self-image as nonprejudiced and that insulate them from recognizing that their behavior is not colorblind.

Much of the research on aversive racism focuses on the orientation of whites toward blacks in the United States, but similar processes are found for the attitudes of members of dominant groups in other countries with strong contemporary egalitarian values and discriminatory histories or policies (e.g., Hodson, Hooper, Dovidio, & Gaertner, 2005). Despite its subtle expression, the consequences of aversive racism are as significant and pernicious as those of the traditional, overt form.

For instance, one of our early experiments (Gaertner & Dovidio, 1977) demonstrates how aversive racism can operate in fairly dramatic ways. In a scenario inspired by an incident in the mid-1960s in which 38 people witnessed the stabbing of a woman, Kitty Genovese, without a single bystander intervening to help, we created a situation in the laboratory in which white participants witnessed a staged emergency involving a black or a white victim. We led some of our participants to believe that they were the only witness to this emergency, while we led others to believe there were other white people who also witnessed the emergency, each isolated in a separate room within the laboratory.

As we predicted, when white participants believed they were the only witness and bore full responsibility for intervening, they helped both white and black victims very frequently (over 85 percent of the time) and equivalently. There was no evidence of blatant racism. In contrast, when they thought other witnesses were in the vicinity and they could rationalize a decision not to help on the basis of some factor other than race, they helped black victims only half as often as they helped white victims

(37.5% versus 75%). This research, therefore, shows that although the bias may be subtle and the people involved may be well-intentioned, its consequences may be severe.

We, and others, find considerable additional empirical support for the aversive racism framework across a range of other paradigms and participant populations. These situations include nonemergency scenarios both inside and outside the laboratory, college admissions and other consequential decisions (see Dovidio & Gaertner, 2004). In addition, in a recent study (Pearson, Dovidio, & Pratto, 2007) we report, as hypothesized within the aversive racism framework, that whereas blatant prejudice is characterized by feelings of antipathy and hate, the subtle bias associated with aversive racism is "cooler" and less affective in nature.

We further propose that because the subtle bias associated with aversive racism occurs without personal awareness and the actions can be attributed, even by observers, to factors other than race, the influence of aversive racism commonly goes unrecognized by whites. As a consequence, whereas blatant expressions of prejudice, such as hate crimes, are readily identified and inhibited by social sanctions, aversive racism is likely to persist relatively unchallenged over time. For instance, Saucier, Miller, and Doucet (2005) report a meta-analysis of 31 experiments conducted over the past 40 years that examined race and whites' helping behavior. Across these studies, Saucier et al. reveal that a pattern of discrimination reflective of aversive racism remained stable over time. Saucier et al. summarize, "the results of this meta-analysis generally supported the predictions for aversive racism theory" (p. 13), and ask, "is racism still a problem in our society? . . . Racism and expressions of discrimination against blacks can and will exist as long as individuals harbor negativity toward blacks at the implicit level" (p. 14).

We propose that prejudice not only systematically influences intergroup *outcomes* but also intergroup *interactions*. In particular, the dissociation between negative implicit attitudes and egalitarian explicit attitudes experienced by aversive racists can have a significant impact on how whites and blacks interact in ways that contribute substantially to misunderstandings in intergroup interactions.

BIAS AND INTERRACIAL INTERACTION

Implicit and explicit attitudes can influence behavior in different ways and under diverse conditions (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995). Explicit attitudes shape deliberative, well-considered responses for which people have the motivation and opportunity to weigh the costs and benefits of various courses of action. Implicit attitudes influence responses that are more difficult to monitor and control. For example, whereas self-reported prejudice predicts overt expressions of bias, measures of implicit attitudes predict

whites' biases in nonverbal behaviors, such as measures of interest (e.g., eye contact), anxiety (e.g., rate of eye blinking), and other cues of friendliness (Dovidio et al., 1997; Fazio et al., 1995). Thus, the relative impact of implicit and explicit attitudes is a function of the situational context, individuals' motivation and opportunity to engage in deliberative processes, and the nature of the behavioral response.

The nature of contemporary racial prejudice in the United States is particularly problematic with respect to producing and perpetuating misunderstandings in interracial interactions. Dovidio, Kawakami, and Gaertner (2002) reveal that whites and blacks have different interpretations of the feelings and actions demonstrated by whites during interracial interactions. Whites have full access to their explicit attitudes and are able to monitor and control their more overt and deliberative behaviors, which are generally expressed in a nonprejudiced and nondiscriminatory way. However, whites do not have full access to their implicit attitudes, which tend to be reflected in their less easily deliberative and controlled behaviors (e.g., nonverbal behaviors).

Perceptions of whites about how they are behaving or how they are perceived by others are based more on their explicit attitudes and overt behaviors, such as the verbal content of their interaction with blacks, and less on their implicit attitudes or less deliberative behaviors. In contrast, the perspective of black interaction partners in these interracial interactions allows them to attend to both the spontaneous (e.g., nonverbal) and deliberative (e.g., verbal) behaviors of whites. To the extent that the black partners attend to whites' nonverbal behaviors, which may signal more negativity than their verbal behaviors, blacks are likely to form more negative impressions of the encounter and be less satisfied with the interaction compared with whites. One fundamental implication of these processes is that whites and blacks are likely to form very different perceptions of race relations, with blacks developing a general sense of distrust of whites (Dovidio, Gaertner, Kawakami, & Hodson, 2002).

IMPLICATIONS OF AVERSIVE RACISM IN LEGAL, EMPLOYMENT, AND HEALTH CARE SETTINGS

Blacks and whites have fundamentally different realities and experiences in the United States. Significant racial disparities exist in legal outcomes, economic opportunity and achievement, and health and health care. In this section, we consider the potential influence of aversive racism in these three fundamental spheres of social life and well-being in the United States. We draw on direct experimental research, our own as well as by others, as well as archival evidence to make a case that aversive racism may play a critical, but largely unrecognized, role in shaping racial disparities in these domains.

Legal Contexts

Traditionally, blacks and whites have not been treated equally under the law (Sidanius & Pratto, 1999). Across time and locations in the United States, blacks have been more likely to be perceived by jurors as guilty, more likely to be convicted of crimes, and, if convicted, sentenced to longer terms for similar crimes, particularly if the victim is white. Although there is some evidence that disparities in judicial outcomes are declining (see Sommers & Ellsworth, 2001), aversive racism may be contributing to the persistence of these disparities.

Three studies represent conceptual replications of the operation of aversive racism in legal decision making over a 25-year period. Faranda and Gaertner (1979) investigated whether being exposed to inadmissible incriminating evidence would have a more negative influence on whites' judgments of black than white defendants. Participants were not exposed to inadmissible evidence (the defendant's confession to a third party), which they were instructed to ignore. This research also tested the hypothesis that this bias could take different forms: Whereas the racial biases of those who are likely to have traditionally racist attitudes (high authoritarians) would reflect primarily antiblack biases, the racial biases of those who are likely to exhibit aversive racism (low authoritarians) would mainly represent pro-white biases.

As expected, both high- and low-authoritarian participants displayed racial bias in their judgments of guilt, but they did so in different ways. High authoritarians were more certain of the black defendant's guilt when they were exposed to the inadmissible evidence than when they were not presented with this testimony. For the white defendant, however, high authoritarians followed the judge's instructions appropriately. Low-authoritarian participants, in contrast, followed the judge's instructions about ignoring the inadmissible testimony when the defendant was black. However, they were biased *in favor of* the white defendant when the inadmissible evidence was presented. Thus, low-authoritarian participants demonstrated a pro–in-group bias.

More than 15 years later, Johnson, Whitestone, Jackson, and Gatto (1995) further examined the impact of the introduction of inadmissible evidence that was damaging to a defendant's case on whites' judgments of a black or white defendant's guilt. No differences in judgments of guilt occurred as a function of defendant's race when all the evidence presented was admissible. However, consistent with the aversive racism framework, the presentation of inadmissible evidence increased judgments of guilt when the defendant was black but not when the defendant was white.

We conducted another study along these lines in the United Kingdom (Hodson et al., 2005), involving inadmissible incriminating DNA evidence. Paralleling the earlier findings, we report that white participants appropriately corrected their judgments for white defendants by effectively

discounting the inadmissible evidence, judging the defendant as less guilty when the damaging evidence was inadmissible than when it was admissible. In contrast, white participants had difficulty suppressing the inadmissible evidence when the defendant was black; they demonstrated a rebound effect, tending to judge the black defendant as more guilty when the evidence was inadmissible than when it was admissible.

Thus, three experiments that used similar paradigms over a span of over 25 years obtained evidence of a subtle but persistent pattern of discrimination predicted by the aversive racism framework. Several other studies of legal decision making have also yielded evidence consistent with the proposition that whites' biases against blacks will be more pronounced when they have an apparently nonrace-related justification for judging a black defendant guilty or sentencing them more severely (Knight, Guiliano, & Sanchez-Ross, 2001). However, also consistent with the aversive racism framework, when testimony is included that racial bias may be involved in the allegations against a black defendant, whites no longer racially discriminate (Sommers & Ellsworth, 2000).

Employment Decisions

The racial gap in median family income has been substantial and enduring over time. The ratio of white family income relative to black family income was 1.73 ($21,904/$12,674) in 1980, 1.72 ($36,915/$21,423) in 1990, 1.57 ($53,029/$33,676) in 2000, and 1.67 ($59,317/35,464) in 2005. Across the same period, the unemployment rate for blacks has hovered around double that of whites: 2.27 (14.3%/6.3%) in 1980, 2.38 (11.4%/4.8%) in 1990, 1.55 (7.6%/4.9%) in 1995, 2.17 (10.4%/4.8%) in 2000, and 1.96 (10.0%/5.1%) in 2005 (U.S. Census Bureau, 2008).

Whereas the evidence indicating the persistence of the effects of aversive racism in whites' helping behavior and juridic decisions have relied on conceptual replications across time and different locations, we have further evidence in employment contexts involving a direct procedural replication (Dovidio & Gaertner, 2000). We used exactly the same paradigm with students from the same college in 1989 and 1999. Participants at these two different points in time were asked to assist in deciding which applicants should be hired as a Resident Advisor in one of the college's large dormitories. They were given background information, including a résumé and excerpts from an interview, indicating (as pretested) that the applicant had very weak, moderate (arguably qualified), or very strong qualifications for the position. Information on the résumé revealed the race of the candidate as black or white. Our prediction, based on the aversive racism framework, was that bias would not be expressed when the candidate was clearly qualified or unqualified for a position, because the appropriate decision would be obvious. However, bias was expected when the appropriate decision was unclear, when the candidate had moderate qualifications.

Consistent with the aversive racism framework, when the candidates' credentials clearly qualified them for the position (strong qualifications) or the credentials clearly were not appropriate (weak qualifications), no discrimination was seen against the black candidate. However, when candidates' qualifications for the position were less obvious and the appropriate decision was more ambiguous (moderate qualifications), white participants recommended the black candidate significantly less often than the white candidate with exactly the same credentials. Moreover, when we compared the responses of participants in 1989 and 1999, whereas overt expressions of prejudice (measured by items on a self-report prejudice scale) declined over this 10-year period, the pattern of subtle discrimination in selection decisions remained essentially unchanged.

More recently, Otero and Dovidio (2005) conceptually replicated this research with human resource professionals in Puerto Rico, focusing on the moderate qualifications and strong qualifications conditions. The findings illustrate the generalizability and persistence of these effects even among experienced professionals in the field. When the applicant had strong qualifications, blacks and whites received equivalently strong recommendations for hiring. However, when the applicant had only moderate qualifications, black candidates were recommended significantly less strongly than white candidates.

Overall, these studies show that, in contrast to the dramatic decline in overt expressions of prejudice, subtle forms of discrimination outlined in the aversive racism framework continue to exist, apparently largely unabated. As we noted earlier, one reason for the persistence of these types of bias is that they are largely shaped by unconscious processes. A recent series of experiments by Son Hing, Chung-Yan, Hamilton, and Zanna (2008) concerning another form of racial bias, prejudice and discrimination against Asian job applicants in Canada, speaks directly to this issue. These researchers also manipulated the strength of qualifications for a white or Asian job candidate. Similar to our findings, more bias was seen against Asian relative to white job applicants when the candidate had moderate qualifications (allowing a nonrace-related excuse for not hiring the person) than when the candidate had very strong credentials. Moreover, Son Hing et al. obtained direct evidence of the role of implicit attitudes. Implicit negative attitudes toward Asians (as measured by an IAT), but not explicit prejudice, predicted weaker support for hiring Asian candidates who had moderate qualifications. However, when the Asian candidate had distinctively strong qualifications (and a failure to hire the applicant was not justifiable on a factor other than race), neither explicit nor implicit prejudice predicted the hiring decision. Support for the Asian applicant was generally strong in this case.

In summary, the research we have presented in this section not only provides further general evidence for the aversive racism framework, but it also illustrates the ways that aversive racism specifically biases

employment decisions. To the extent that current antidiscrimination laws in the United States require clear evidence that race was *the* determining factor (not some other factor) and bias was intentional, the subtle effect of aversive racism may be relatively immune to prosecution. As shown in our replications over time, the effects of aversive racism in employment contexts may occur largely unrecognized, and as a consequence remain largely unaddressed, over time.

Health Care

Significant and persistent disparities occur in the quality of care received by black patients compared to white patients in the United States (Geiger, 2003; Smedley, Stith, & Nelson, 2003). With respect to care for coronary heart disease, for example, blacks are less likely than whites to be seen by a coronary specialist, less likely to be prescribed appropriate preventive and emergency medications for their heart disease, and less likely to receive surgical procedures intended to remedy various types of coronary heart disease (see Dovidio et al., 2008, for a review).

Despite ample evidence of racial disparities in health and health care, controlled experimental studies investigating subtle discrimination among health care professionals are rare in the literature. Thus, in this section we consider the *implications of* the aversive racism framework for understanding health care disparities.

Consistent with the research on aversive racism in other domains, which reveals that discriminatory actions are more likely to occur when situational demands are unclear or when norms for appropriate actions are weak or ambiguous (Dovidio & Gaertner, 2004), racial disparities in treatment are greater when physicians engage in "high-discretion" procedures, such as recommending a test or making a referral for a procedure or drug, than in "low-discretion" procedures, such as emergency surgery (Geiger, 2003). For example, racial disparities in prostate cancer, for which physicians generally have considerable discretion with regard to screening and treatment, are particularly pronounced. Black men are significantly less likely than white men to receive screening tests for prostate cancer. In addition, physicians are more likely to delay active treatment of blacks than whites for prostate cancer (see Dovidio et al., 2008, for a review).

Whereas most studies in the coronary heart disease literature have used archival data, one notable exception is a study by Schulman et al. (1999) involving primary care physicians at a national conference. This experiment reveals the causal influence of patient race on medical decisions. Physicians viewed videotapes of an actor playing the role of a patient complaining about chest pain, and gender and race of the patient (black or white) were systematically manipulated. Women and blacks were significantly less likely to be referred for further testing than were men and whites, and black women were only 40 percent as likely to receive such

a referral as were white men. Although this experimental study has been challenged methodologically, the results are consistent with the findings from archival data.

Recently, Green et al. (2007) extend this line of research to explore the role in which physicians' explicit and implicit racial attitudes and stereotypes might affect their medical decisions. Their findings support the basic premise of the aversive racism framework that people who appear nonprejudiced, and presumably believe they are egalitarian, may be influenced by their unconscious biases to discriminate against blacks in subtle but significant ways. In this study, physicians reported no explicit biases toward blacks relative to whites, but had more negative implicit attitudes. Moreover, the more negative these implicit attitudes were, the less likely respondents were to recommend the thrombolytic drugs for black patients.

In summary, consistent with the aversive racism framework, racial bias appears to have a systematic influence in health care delivery, particularly in high discretion contexts. This bias is also related to providers' implicit racial biases. Moreover, research on aversive racism suggests that because racial bias is subtle and can be rationalized, its operation and impact are typically underestimated among whites. Indeed, Epstein (2005) argues that the Institute of Medicine (IOM) Panel on Racial and Ethnic Disparities in Health Care (Smedley et al., 2003) overemphasized the potential role of the racial biases in accounting for disparities in health care. Epstein stated, "it is doubtful that hidden forms of discrimination are prevalent in a profession whose professional norms are set so strongly against it" (p. 26). However, as the research on aversive racism demonstrates, genuinely strong egalitarian values do not insulate people against being subtly biased.

COMBATING AVERSIVE RACISM

The social psychological literature can also help guide the development of practical interventions that can attenuate and reduce racial bias. Traditional prejudice-reduction techniques have been concerned with changing conscious attitudes—old-fashioned racism—and obvious expressions of bias, and have utilized traditional educational programs or persuasive messages to change explicit attitudes (Stephan & Stephan, 2001). However, because of its pervasiveness, subtlety, and complexity, the traditional techniques for eliminating racial bias are ineffective for combating aversive racism. Aversive racists already recognize prejudice as harmful, but they do not recognize that *they* are prejudiced. Other techniques are thus required.

Addressing Unconscious Attitudes and Beliefs

As we described earlier, aversive racism is characterized by conscious (explicit) egalitarian attitudes and negative unconscious (implicit) attitudes and beliefs. Simply because implicit attitudes are unconscious and

automatically activated, however, does not mean they are immutable to change. To the extent that unconscious attitudes and stereotypes are associations learned through socialization, we propose that they can also be unlearned or inhibited by equally well-learned countervailing influences. We posit that with extensive practice, it is possible to change even implicit beliefs. For example, extended practice in associating counter-stereotypic characteristics with a group can inhibit or suppress the "automatic" activation of cultural stereotypes (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000).

The problem in practice is that whites are typically motivated to avoid seeing themselves are racially biased and often try to behave as such. Whites often adopt a colorblind orientation, particularly when they anticipate racial tension. However, efforts to be colorblind and suppress acknowledgement of race can produce a "rebound effect," in which implicit attitudes become activated even more. Furthermore, because minorities seek acknowledgement of their racial identity, attempts by whites to be colorblind may alienate minority group members and further contribute to racial distrust (Dovidio et al., 2002).

In efforts to reduce prejudice, it may be possible to capitalize on aversive racists' good intentions and induce self-motivated efforts to reduce unconscious biases by making them aware. Work by Monteith and Voils (1998) indicates that when low-prejudiced people recognize discrepancies between their behavior (i.e., what they *would* do) and their personal standards (i.e., what they *should* do) toward minorities, they feel guilt and compunction, which subsequently produces motivations to respond without prejudice in the future. With practice over time, these individuals learn to reduce prejudicial responses and respond in ways that are consistent with their nonprejudiced personal standards. This process of self-regulation may produce changes in even unconscious negative responses when extended over time (Dovidio, Kawakami, & Gaertner, 2000).

Redirecting In-group Bias

One basic argument we have made in our analysis of social biases is that the negative feelings that develop toward other groups may be rooted, in part, in fundamental, normal psychological processes. One such process is the categorization of people into in-groups and out-groups. As we noted earlier, social categorization contributes to aversive racism. Because categorization is a basic process fundamental to intergroup bias, we have targeted this process as an avenue through which we may attempt to attenuate and reduce the negative repercussions of aversive racism. To do this we have proposed the Common Ingroup Identity Model (Gaertner & Dovidio, 2000).

The Common Ingroup Identity Model is rooted in the social categorization perspective of intergroup behavior and recognizes the central role of social categorization in intergroup bias. Specifically, if members of

different groups are induced to think of themselves as a single superordinate group rather than as two separate groups, attitudes toward former out-group members will become more positive through in-group bias. Thus, by changing the basis of categorization from race to an alternative dimension, one can alter who "we" is and who "they" are, undermining a contributing force to contemporary racism. Formation of a common identity, however, does not necessarily require groups to forsake their other identities. It is possible for members to conceive of themselves as holding a "dual identity" in which other identities and the superordinate group identity are salient simultaneously.

CONCLUSION

The influence of aversive racism is pervasive; it affects race relations and outcomes for blacks in a variety of ways. It persists because it remains largely unrecognized and thus unaddressed. However, we propose that it can be combated with new approaches and strategies that are uniquely targeted at critical components of aversive racism. For example, because aversive racists are truly motivated to be nonprejudiced, making them aware, in a nonthreatening way, of their unconscious biases can arouse motivations to change in fundamental ways, which can eventually reduce unconscious biases. Nevertheless, without sufficient recognition of the subtle nature of contemporary biases and without the appropriate tools for combating these particular biases, significant progress toward a truly just society will be difficult to achieve. Good intentions alone are not good enough.

NOTE

Preparation of this chapter was supported by NSF Grant # BCS-0613218 awarded to the first two authors.

REFERENCES

Blair, I. V. (2001). Implicit stereotypes and prejudice. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the Legacy and Future of Social Cognition* (pp. 359–374). Mahwah, NJ: Erlbaum.

Bobo, L. (2001). Racial attitudes and relations at the close of the twentieth century. In N. J. Smelser, W. J. Wilson, & F. Mitchell (Eds.), *Racial trends and their consequences* (Vol. 1, pp. 264–301). Washington, DC: National Academy Press.

Dovidio, J. F. (2001). On the nature of contemporary prejudice: The third wave. *Journal of Social Issues, 57*, 829–849.

Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science, 11*, 319–323.

Dovidio, J. F., & Gaertner, S. L. (2004). Aversive racism. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (vol. 36, pp. 1–51). San Diego, CA: Academic Press.

Dovidio, J. F., Gaertner, S. L., Kawakami, K., & Hodson, G. (2002). Why can't we just get along? Interpersonal biases and interracial distrust. *Cultural Diversity & Ethnic Minority Psychology, 8,* 88–102.

Dovidio, J., Kawakami, K., & Beach, K. (2001). Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias. In R. Brown & S. L. Gaertner (Eds.), *Blackwell handbook of social psychology* (Vol. 4, pp. 175–197). Oxford, UK: Blackwell.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2000). Reducing contemporary prejudice: Combating explicit and implicit bias at the individual and intergroup level. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 137–163). Hillsdale, NJ: Erlbaum.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology, 82,* 62–68.

Dovidio, J., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). The nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33,* 510–540.

Dovidio, J. F., Penner, L. A., Albrecht, T. L., Norton, W. E., Gaertner, S. L., & Shelton, J. N. (2008). Disparities and distrust: The implications of psychological processes for understanding racial disparities in health and health care. *Social Science and Medicine, 67,* 476–487.

DuBois, W.E.B. (1903/1986). *Writings* (reprinted, W. E. Burghardt, Ed.). New York: Viking.

Epstein, R. A. (2005). Disparities and discrimination in health care coverage: A critique of the Institute of Medicine study. *Perspectives in Biology and Medicine, 48*(1, Suppl.), S26–S41.

Faranda, J., & Gaertner, S. L. (1979, March). *The effects of inadmissible evidence introduced by the prosecution and the defense, and the defendant's race on the verdicts by high and low authoritarians.* Paper presented at the annual meeting of the Eastern Psychological Association, New York.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A *bona fide* pipeline? *Journal of Personality and Social Psychology, 69,* 1013–1027.

Gaertner, S. L., & Dovidio, J. F. (1977). The subtlety of White racism, arousal, and helping behavior. *Journal of Personality and Social Psychology, 35,* 691–707.

Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). Orlando, FL: Academic Press.

Gaertner, S. L., & Dovidio, J. F. (2000). *Reducing intergroup bias: The Common Ingroup Identity Model.* Philadelphia, PA: The Psychology Press.

Geiger, H. J. (2003). Racial and ethnic disparities in diagnosis and treatment: A review of the evidence and a consideration of causes. In B. Smedley, A. Stith, & A. Nelson (Eds.), *Unequal Treatment: Confronting racial and ethnic disparities in health care* (pp. 417–454). Washington, DC: The National Academies Press.

Green, A. R., Carney, D. R., Pallin, D. J., Ngo, L. H., Raymond, K. L., Iezzoni, L. I., et al. (2007). Implicit bias among physicians and its predictions of thrombolysis decisions for Black and White patients. *Journal of General Internal Medicine, 22,* 1231–1238.

Greenwald, A., McGhee, D., & Schwartz, J. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464–1480.

Hodson, G., Hooper, H., Dovidio, J. F., & Gaertner, S. L. (2005). Aversive racism in Britain: Legal decisions and the use of inadmissible evidence. *European Journal of Social Psychology, 35,* 437–448.

Johnson, J. D., Whitestone, E., Jackson, L. A., & Gatto, L. (1995). Justice is still not colorblind: Differential racial effects of exposure to inadmissible evidence. *Personality and Social Psychology Bulletin, 21,* 893–898.

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in trait negation on stereotype activation. *Journal of Personality and Social Psychology, 78,* 871–888.

Knight, J. L., Guiliano, T. A., & Sanchez-Ross, M. G. (2001). Famous or infamous? The influence of celebrity status and race on perceptions of responsibility for rape. *Basic and Applied Social Psychology, 23,* 183–190.

Kovel, J. (1970). *White racism: A psychohistory.* New York: Pantheon.

Monteith, M. J., & Voils, C. (1998). Proneness to prejudiced responses: Toward understanding the authenticity of self-reported discrepancies. *Journal of Personality and Social Psychology, 75,* 901–916.

Otero, L., & Dovidio, J. F. (2005). Unpublished data. Department of Psychology, University of Connecticut, Storrs, CT.

Pearson, A. R., Dovidio, J. F., & Pratto, F. (2007). Racial prejudice, intergroup hate, and blatant and subtle bias of Whites toward Blacks in legal decision making in the United States. *International Journal of Psychology and Psychological Therapy, 7,* 145–158.

Saucier, D. A., Miller, C. T., & Doucet, N. (2005). Differences in helping Whites and Blacks: A meta-analysis. *Personality and Social Psychology Review, 9,* 2–16.

Schulman, K. A., Berlin, J. A., Harless, W., Kerner, J. F., Sistrunk, S., Gersh, B. J., et al. (1999). The effect of race and sex on physicians' recommendations for cardiac catheterization. *The New England Journal of Medicine, 340*(8), 618–626.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression.* New York: Cambridge University Press.

Smedley, B. D., Stith, A. Y., & Nelson, A. R. (Eds.). (2003). *Unequal treatment: Confronting racial and ethnic disparities in health care.* Washington, DC: National Academy Press.

Sommers, S. R., & Ellsworth, P. C. (2000). Race in the courtroom: Perceptions of guilt and dispositional attributions. *Personality and Social Psychology Bulletin, 26,* 1367–1379.

Sommers, S. R., & Ellsworth, P. C. (2001). White juror bias: An investigation of prejudice against Black defendants in the American courtroom. *Psychology, Public Policy, and Law, 7,* 201–229.

Son Hing, L. S., Chung-Yan, G., Hamilton, L., & Zanna, M. (2008). A two-dimensional model that employs explicit and implicit attitudes to characterize prejudice. *Journal of Personality and Social Psychology, 94,* 67–81.

Stephan, W. G., & Stephan, C. W. (2001). *Improving intergroup relations.* Thousand Oaks, CA: Sage.

U.S. Census Bureau. (2008). *The 2008 statistical abstract.* Retrieved May 15, 2008, from http://www.census.gov/compendia/statab/cats/labor_force_employment_earnings.html.