AVERSIVE RACISM

John F. Dovidio

Samuel L. Gaertner

Race relations in the United States, especially in terms of whites' orientations toward blacks, have been characterized by inconsistencies and ambivalence. This duality has existed virtually from the beginning. The United States was founded on the principle of "liberty and justice for all" and on "the proposition that all men are created equal." Nevertheless, it was not until 175 years after these basic human rights were proclaimed by the Declaration of Independence and guaranteed by the Constitution that the initial civil rights legislation was passed and the United States formally recognized that black and white people were equal under the law. Such structural changes in the law, however, do not mean that racial attitudes of white Americans will quickly reflect this new legal standard.

Myrdal (1944) identified the paradox between historical egalitarian values and racist traditions in the United States, describing the "American dilemma." According to Myrdal, the dilemma involves

> The ever-raging conflict between, on the one hand, the valuations preserved on the general plane which we call the "American creed," where the American thinks, talks, and acts under the influence of high national and Christian precepts and, on the other hand, the valuations on the specific planes of individual and group living, where personal and local interests; economic, social, and sexual jealousies; consideration of community prestige and conformity; group prejudice against particular persons or types of people; and all sorts of miscellaneous wants, impulses, and habits dominate his outlook. (p. xliii)

The American dilemma of 1944 was thus between egalitarian ideals and personal and social forces that promote prejudice and discrimination. In this chapter we examine one contemporary legacy of this historical dilemma—"aversive racism."

Elements of the American dilemma still remain. On the one hand, in the United States the principle of equality continues as a fundamental social value (Schuman, Steeh, Bobo, & Krysan, 1997). Moreover, whites' expressions of prejudice toward traditionally underrepresented groups, and toward blacks in particular, have declined substantially over time. As Bobo (2001) concluded 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). From 1960 to the present, public opinion polls have revealed that whites increasingly support integration in schools, public transportation, jobs, and housing; whites' support for interracial marriage has also grown correspondingly.

On the other hand, evidence of racial disparity and discrimination still remain. Although disparities can have causes other than discrimination, economic indices show consistent differences in status based on race. For instance, the median family income for blacks is less than two-thirds that of whites, a differential that has widened over the past two decades (Blank, 2001). Also, on several basic measures of health and well-being, the racial gap has been maintained or, in some cases (e.g., infant mortality), has widened substantially over the past 50 years (Jenkins, 2001). Furthermore, recent studies suggest that over the life span black and white patients receive unequal treatment from medical practitioners resulting in less favorable health-related outcomes for blacks (see Smedley, Stith, & Nelson, 2003; Whaley, 1998). Steady trends toward residential integration that were observed from the 1950 to 1970 have slowed in the South and stagnated in the North (Massey, 2001). Massey (2001) observed, "Either in absolute terms or in comparison to other groups, blacks remain a very residentially segregated and spatially isolated people" (p. 403).

Evidence suggests that discrimination is a key factor in many of these disparities. In terms of career dynamics, whites have an advantage over blacks in initial wage level and opportunities for training (Rosenfeld, 1998). In addition, blacks are more likely to be discriminated against relative to whites when *economic* conditions require layoffs (Elvira & Zatzick, 2002, see also Murphy-Berman, Berman, & Campbell, 1998). In addition, differences in health care, which have been attributed to racism by health providers (Smedley et al., 2003), occur over and above access to health insurance.

Clearly, discrimination and perceptions of discrimination continue to be dominant forces in the lives of minorities in the United States. For example,

within the government workforce, 55% of blacks and 28% of Hispanics reported that discrimination hinders their career advancement (U.S. Merit Systems Protection Board, 1997). In the general public, nearly half of black Americans (47%) reported that they were treated unfairly during the previous month in their own community in at least one of five common situations: while shopping, at work, in restaurants or other entertainment places, in dealing with the police, and using public transportation (Gallup, 2002). Moreover, white and black Americans have very different perceptions of the racial discrimination that blacks face today. More than three fourths (79%) of whites reported that blacks "have as good a chance as whites" to "get any kind of job," but fewer than half (46%) of blacks shared that view. Whereas the vast majority (69%) of whites perceived that blacks were treated "the same as whites," the majority of blacks (59%) reported that blacks were treated worse than whites.

Over the past 35 years, we have explored the nature of whites' racial attitudes to understand this duality between the generally expressed non-prejudicial views of whites in contemporary U.S. society and the persistence of significant racial disparity and discrimination. Our work built on the conceptual framework of Kovel (1970), who distinguished between dominative and aversive racism. Dominative racism is the "old-fashioned," blatant form. According to Kovel, the dominative racist is the "type who acts out bigoted beliefs—he represents the open flame of racial hatred" (p. 54). Aversive racists, in comparison, sympathize with victims of past injustice, support the principle of racial equality, and regard themselves as nonprejudiced, but, at the same time, possess negative feelings and beliefs about blacks, which may be unconscious. Aversive racism is hypothesized to be qualitatively different than blatant, "old-fashioned," racism, is more indirect and subtle, and is presumed to characterize the racial attitudes of most well-educated and liberal whites in the United States. Nevertheless, the consequences of aversive racism (e.g., the restriction of economic opportunity) are as significant and pernicious as those of the traditional, overt form (Dovidio & Gaertner, 1998; Gaertner & Dovidio, 1986).

This chapter reviews this program of research, highlighting basic assumptions, key findings, and major developments. Although our research has focused on race relations in the United States, the processes of aversive racism are not limited by national or geographic boundaries and could reflect attitudes toward a number of different groups when overt forms of discrimination are recognized as inappropriate (see Esses, Dovidio, Jackson, & Armstrong, 2001; Kleinpenning & Hagendoorn, 1993; Pettigrew & Meertens, 1995). We begin by reviewing the nature of aversive racism, including the contributing psychological factors and the potential conflict between whites' conscious endorsement of egalitarian principles and unconscious negative

feelings and beliefs about blacks. We next consider basic evidence about how aversive racism operates and its moderating factors. Then we consider our research examining separately the conscious and unconscious components of aversive racism. In the last two sections we explore ways of combating the effects of aversive racism and consider conclusions and implications.

I. The Nature of Aversive Racism

A critical aspect of the aversive racism framework is the conflict between whites' denial of personal prejudice and underlying unconscious negative feelings toward and beliefs about blacks. Because of current cultural values, most whites have strong convictions concerning fairness, justice, and racial equality. However, because of a range of normal cognitive, motivational, and sociocultural processes that promote intergroup biases, most whites also develop some negative feelings toward or beliefs about blacks, of which they are unaware or which they try to dissociate from their nonprejudiced self-images. These negative feelings that aversive racists have toward blacks do not reflect open hostility or hatred. Instead, aversive racists' reactions may involve discomfort, uneasiness, disgust, and sometimes fear. That is, they find blacks "aversive," while at the same time find any suggestion that they might be prejudiced "aversive" as well. Thus, aversive racism may involve *more positive reactions* to whites than to blacks, reflecting a pro–in-group rather than an anti–out-group orientation, thereby avoiding the stigma of overt bigotry and protecting a nonprejudiced self-image.

The existence of these nearly unavoidable racial biases and the simultaneous desire to be nonprejudiced represents a basic duality of attitudes and beliefs for aversive racists that can produce racial ambivalence (see also Katz & Hass, 1988; Katz, Wackenhut, & Hass, 1986). We recognize that all racists are not aversive or subtle, that old-fashioned racism still exists, that there are individual differences in aversive racism, and that some whites may not be racist at all. Nevertheless, we propose that aversive racism generally characterizes the racial attitudes of a large proportion of whites who express apparently nonprejudiced views. Some of the sources of this negative affect and beliefs are reviewed in the next section.

A. NEGATIVE ATTITUDES AND RACIAL AMBIVALENCE

In contrast to traditional approaches that emphasize the *psychopathology* of prejudice (e.g., Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950; see Duckitt, 1992), the negative feelings and beliefs that underlie

aversive racism are hypothesized to be rooted *in normal*, often adaptive, psychological processes.

These processes fundamentally involve the consequences of social categorization. People inherently categorize others into groups, typically in ways that delineate one's own group from other groups (Hamilton & Trolier, 1986). The mere categorization of people into in-groups and out-groups, even on the basis of arbitrary assignment, is sufficient to initiate (often spontaneously; Otten & Moskowitz, 2000) an overall evaluative bias in which people categorized as members of one's own group are evaluated more favorably than are those perceived as members of another group (Brewer, 1979; Tajfel, 1970).

Social categorization also fundamentally influences how people process information about others. Perceptually, when people or objects are categorized into groups, actual differences between members of the same category tend to be minimized (Tajfel, 1969) and often ignored in making decisions or forming impressions, whereas between-group differences tend to become exaggerated (Abrams, 1985; Turner, 1985). Cognitively, people retain more information in a more detailed fashion for ingroup members than for outgroup members (Park & Rothbart, 1982), have better memory for information about ways in which ingroup members are similar to and outgroup members are dissimilar to the self (Wilder, 1981), and remember less positive information about outgroup members (Howard & Rothbart, 1980). In addition, they evaluate the products of ingroup members more favorably than those of outgroup members (Ferguson & Kelley, 1964) and they work harder for groups identified as ingroups (Worchel, Rothgerber, Day, Hart, & Butemeyer, 1998). The process of social categorization also influences affective reactions. As Insko et al. (2001) have demonstrated, categorization in terms of group membership rather than individual identity evokes greater feelings of fear and lower levels of trust in interactions with others.

In the United States, because of its historical and central importance, social categorization by race is virtually automatic. Without effort or control, whites spontaneously differentiate people by race, with the activation of racial categories being triggered by the actual or symbolic presence of a black person. Because of sociocultural influences, these racial categories are associated with traditional negative stereotypes of blacks (Devine, 1989; Dovidio, Evans, & Tyler, 1986), as well as negative attitudes (Dovidio, Kawakami, & Beach, 2001).

Motivational factors, such as needs for power, status, and control (Operario & Fiske, 1998), also contribute to whites' biased feelings and beliefs. Although these needs may be personal, they are also often group related. In Social Identity Theory, Tajfel and Turner (1979) proposed that a

person's need for positive identity may be satisfied by membership in prestigious social groups. This need also motivates social comparisons that favorably differentiate ingroup from outgroup members. Discrimination is one way of creating "positive distinctiveness" for one's group, which, in turn, can boost one's self-esteem and promote feelings of control and superiority (Fein & Spencer, 1997; Tajfel & Turner, 1979).

Competition and perceived competition between groups further exacerbate intergroup biases (Campbell, 1965; Esses et al., 2001; Sherif, 1966). These biases are functional. Tangibly, discrimination offers economic advantages to members of the majority group and maintains that group's political, social, and corporate power (Blumer, 1958; Bobo, 1999). Because blacks have traditionally been perceived to threaten whites' basic values and well-being, powerful cognitive and motivational forces that are a function of social categorization and perceived threat form the basis for the negative racial feelings of aversive racists.

Because of current cultural values, however, most whites also have convictions concerning fairness, justice, and racial equality (Bobo, 2001). At the same time, however, psychological and social forces contribute to whites' egalitarian orientation. Principles of fairness, justice, and equity are universal, and they profoundly shape human functioning and social life (Kelman, 2001). Equality and justice are not only fundamental principles in the United States, but they are steadily endorsed more strongly and broadly over time and by succeeding generations (Schuman et al., 1997). The vast majority of white Americans today believe that prejudice and discrimination are wrong, and they indicate strong support for social and political equality (Bobo, 2001). Thus, it is the existence of both almost unavoidable racial biases and conscious adherence to nondiscriminatory principles that forms the basis of the ambivalence that aversive racists experience.

We note that other forms of contemporary racial biases, such as *modern racism* (McConahay, 1986) and *symbolic racism* (Sears, Henry, & Kosterman, 2000) also hypothesize a conflict between the denial of personal prejudice and underlying unconscious negative feelings and beliefs. What distinguishes aversive racism from modern and symbolic racism is the nature of the conscious beliefs that permit discrimination to be expressed. Symbolic Racism Theory emphasizes that beliefs about individualism and meritocracy that become racialized motivate opposition to policies designed to benefit racial and ethnic minorities. For instance, a measure of individual differences in symbolic racism predicted support for Proposition 209, which was designed to dismantle affirmative action in California in 1996 (Sawires & Peacock, 2000). Modern Racism Theory similarly proposes that beliefs associated with conservative ideologies can justify discriminatory behaviors, but this theory places more emphasis on the moderating effects

of contexts that provide a justification for negative responses to minorities (e.g., a previous negative decision about a comparable white job candidate; McConahay, 1983).

Whereas modern and symbolic racism characterize the attitudes of political conservatives, the aversive racism framework focuses on biases among people who are politically liberal and who openly endorse nonprejudiced beliefs, but whose *unconscious* negative feelings get expressed in subtle, indirect, and rationalizable ways (Gaertner & Dovidio, 1986). For example, research by Nail, Harton, and Decker (2003, Studies 1 and 2) revealed that in a context in which race was very salient, politically liberal participants responded more positively toward a black than toward a white person, a pattern that could be expected among aversive racists, whereas politically conservative respondents responded more favorably toward the white person. In further support of the aversive racism framework, Nail et al. (Study 3) found that only liberals displayed greater physiological arousal to the touch of a black versus a white person, which Nail et al. argued reflected the intrapsychic conflict associated with aversive racism.

II. The Operation of Aversive Racism

The aversive racism framework also helps to identify when discrimination against blacks and other minority groups will or will not occur. Whereas old-fashioned racists exhibit a direct and overt pattern of discrimination, aversive racists' actions may appear more variable and inconsistent. Sometimes they discriminate (manifesting their negative feelings), and sometimes they do not (reflecting their egalitarian beliefs). Our research has provided a framework for understanding this pattern of discrimination.

Because aversive racists consciously recognize and endorse egalitarian values and because they truly aspire to be nonprejudiced, they will *not* discriminate in situations with strong social norms when discrimination would be obvious to others and to themselves. Specifically, we propose that when people are presented with a situation in which the normatively appropriate response is clear, in which right and wrong is 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. To avoid the attribution of racist intent, aversive racists either will treat blacks and whites equally, or they will respond even more favorably to blacks than to whites. Wrongdoing, which would directly threaten their nonprejudiced self-image, would be too costly. However, because aversive racists still possess feelings of uneasiness,

these feelings will eventually be expressed, but they will be expressed 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 color blind.

Generally, then, aversive racists may be identified by a constellation of characteristic responses to racial issues and interracial situations. First, aversive racists, in contrast to old-fashioned racists, endorse fair and just treatment of all groups. Second, despite their conscious good intentions, aversive racists unconsciously harbor feelings of uneasiness toward blacks and thus try to avoid interracial interaction. Third, when interracial interaction is unavoidable, aversive racists experience anxiety and discomfort, and consequently they try to disengage from the interaction as quickly as possible. Fourth, because part of the discomfort that aversive racists experience is due to a concern about acting inappropriately and appearing prejudiced, aversive racists strictly adhere to established rules and codes of behavior in interracial situations that they cannot avoid. Finally, their feelings will be expressed, but in subtle, unintentional, rationalizable ways that disadvantage minorities or unfairly benefit the majority group. Nevertheless, in terms of conscious intent, aversive racists intend not to discriminate against people of color—and they behave accordingly when it is possible for them to monitor the appropriateness of their behavior.

The aversive racism framework thus considers the situation as a critical factor influencing the expression of racial bias by whites toward blacks. Although social influences can directly affect the level of bias that is expressed (Pettigrew, 1959), we emphasize the *moderating* role of situational factors on whether the unconscious negative aspects of aversive racists' attitudes are manifested in terms of racial discrimination. That is, whether the situation is one in which a negative act toward a black person would be attributed to racial intent, by others or by the aversive racist himself or herself, determines whether bias will be expressed.

We have found consistent support for the aversive racism framework across a broad range of situations (see Dovidio & Gaertner, 1998; Gaertner & Dovidio, 1986). Much of the research reported in this chapter focuses on the responses of white college students—well-educated and typically liberal people—who are presumed to represent a prime population for aversive racism. Nevertheless, we note that many of the findings and principles we

discuss extend to biases exhibited by liberal noncollege populations (e.g., Gaertner, 1973). In the next sections, we describe examples of a series of different studies to illustrate the operation of aversive racism. The evidence presented in this section is derived from a variety of different paradigms, such as ones involving interventions to help people in need and employment or admission selection decisions.

A. SERENDIPITY AND AVERSIVE RACISM

We began our research on racism naively with a simple assumption: based on differences in their expressed racial attitudes (see Adorno et al., 1950), conservative whites would behave in a more racially discriminatory way than would liberal whites. However, we discovered, somewhat serendipitously, that racial discrimination was complex, and it occurs in subtle as well as overt ways.

In an initial study of contemporary racism and interracial helping (Gaertner, 1973), white participants residing in Brooklyn, New York, were selected for a field experiment on helping on the basis of their liberal or conservative orientations, as indicated by their political party affiliations (i.e., liberal or conservative parties in New York State) that were a matter of public record. Both the liberal and the conservative households received ostensibly wrong-number telephone calls that quickly developed into requests for assistance. The callers, who were clearly identifiable from their dialects as being black or white, explained that their car was disabled and that they were attempting to reach a service garage from a public telephone along the parkway. The callers further claimed that they had no more change to make another call and asked the participant to help by calling the garage. If the participant agreed to help and called the number, ostensibly of the garage, a "helping" response was scored. If the participant refused to help or hung up after the caller explained that he or she had no more change, a "not helping" response was recorded. If the participant hung up before learning that the motorist had no more change, the response was recorded as a "premature hang-up."

The first finding from this study was direct and predicted. Conservatives showed a higher "helping" response to whites than to blacks (92% vs. 65%), whereas liberals helped whites somewhat, but not significantly, more than blacks (85% vs. 75%). By this measure, conservatives were more biased against blacks than were liberals. Additional inspection of the data, however, revealed an unanticipated finding. Liberals "hung up prematurely" much more often on blacks than they did on whites (19% vs. 3%) and especially often on a black male motorist (28%). Conservatives did not discriminate in

this way (8% vs. 5%). From the perspective of black callers, the consequence of a direct "not helping" response and of a "premature hang-up" was the same: they would be left without assistance. From the perspective of the participants, however, the consequences were different. Whereas a "not helping" response was a direct, intentional form of discrimination because it should have been clear to participants that their help was needed, a "premature hang-up" was a more indirect form because participants disengaged from the situation before they learned of the other person's dependence on them, and thus participants never overtly refused assistance. Indeed, to refuse help that is perceived to be needed clearly violates the social responsibility norm, whereas the appropriateness of hanging up prematurely is unclear. Therefore, both conservative and liberal whites discriminated against blacks but in different ways.

III. Emergency Intervention

Another of our early experiments (Gaertner & Dovidio, 1977) demonstrates how aversive racism can operate in fairly dramatic ways. The scenario for the experiment was 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. What accounted for this behavior? Feelings of responsibility play a key role (see Darley & Latané, 1968). If a person witnesses an emergency knowing that he or she is the only bystander, that person bears all of the responsibility for helping and, consequently, the likelihood of helping is high. In contrast, if a person witnesses an emergency but believes that there are several other witnesses who might help, then the responsibility for helping is shared. Moreover, if the person believes that someone else will help or has already helped, the likelihood of that bystander taking action is significantly reduced.

We created a situation in the laboratory in which white participants witnessed a staged emergency involving a black or white victim. We led some of our participants to believe that they would be the only witness to this emergency, whereas we led others to believe that there would be other white people who also witnessed the emergency. We predicted that, because aversive racists do not act in overtly bigoted ways, whites would not discriminate when they were the only witness and the responsibility for helping was clearly focused on them. However, we anticipated that whites would be much less helpful to black than to white victims when they had a justifiable excuse not to get involved, such as the belief that one of the *other* witnesses would take responsibility for helping.

The results clearly reflected these predictions. When white participants believed that they were the only witness, they helped both white and black victims very frequently (more than 85% of the time) and equivalently. There was no evidence of blatant racism. In contrast, when they thought there were other witnesses 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 white victims (37.5% vs. 75%). Thus, these results illustrate the operation of subtle biases in relatively dramatic, spontaneous, and life-threatening circumstances involving a failure to help, rather than an action intentionally aimed at doing harm. Therefore, this research shows that although the bias may be subtle and the people involved may be well intentioned, its consequences may be severe.

Across a range of other studies using a number of helping paradigms, we found evidence that discrimination by whites against blacks occurs primarily when norms for appropriate behavior are weak or ambiguous (Frey & Gaertner, 1986) and tend to be more pronounced when the interaction involves potential threats to the traditionally superior status of whites relative to blacks (Dovidio & Gaertner, 1981, 1983a).

Another series of studies provides evidence of the potential influence of aversive racism on another type of helping (Turner & Pratkanis, 1994), support for public policies designed to benefit blacks.

A. POLICY SUPPORT

Affirmative action has been one of the most hotly debated policies in American politics over the past three decades, virtually since the inception of the program (Skrentny, 1996). As the recent *Grutter v. Bollinger et al.* Supreme Court decision demonstrates (Supreme Court of the United States, 539 U.S., 2003), the issues remain contentious today.

One popular criticism of affirmative action centers on negative reactions based on the perceived unfairness of these policies. The protest by many whites expressed around the *Regents of the University of California v Bakke* (1978) case was that the medical school admissions procedures were a form of "reverse discrimination" that violated their fundamental beliefs about procedural justice and fairness. That is, the commonly articulated reason for challenging the admissions procedure was that the policy was discriminatory and negated individual selection, evaluation, and advancement based on merit.

Considerable theoretical and empirical support exists for the idea that procedural fairness is a critical factor in determining people's response to decision-making procedures (Lind, Kurtz, Musante, Walker, & Thibaut, 1980). With respect to affirmative action, the more weight given to

category-based criteria, such as race or sex, the less fair the procedure is perceived to be and the more negative the reactions to the policy and the persons involved. In addition, consistent with this reasoning, people who are more committed to principles of merit tend to be more opposed to affirmative action when they believe that discrimination is no longer a problem but tend to be more supportive of affirmative action when they recognize the persistent effects of discrimination (Son Hing, Bobocel, & Zanna, 2002).

In two studies, we therefore examined the importance of how affirmative action is framed on whites' reactions to affirmative action. Whereas many people may initially feel that affirmative action policies are unfair when characteristics such as race or sex are weighed in the decision at the micro level, they may come to perceive the procedure as fair if, at a macro level, they recognize the value of diversity—"that individuals bring with them into the organization not merely different amounts of the same things, but also different kinds of things that make them valuable to an organization" (Clayton & Tangri, 1989, p. 180). Similarly, although preferential action may be seen as unfair in a specific case, the same action may be perceived as more fair if it is presented as a compensatory response to address historical inequities.

Racism may also be a factor in attitudes toward affirmative action. In general, people higher in self-reported racial prejudice are more opposed to affirmative action (Frederico & Sidanius, 2002; Kravitz, 1995). We have proposed that aversive racism also plays a significant role in opposition to affirmative action policies, particularly those designed to support blacks.

In one of our studies that addressed the influence of aversive racism on opposition to affirmative action (Murrell, Dietz-Uhler, Dovidio, Gaertner, & Drout, 1994), white respondents were questioned about their perceptions of fairness and support for four common ways of presenting affirmative action policies. Two of these policies focused on micro-level actions varying in the degree to which the action places emphasis on nonmerit factors to address disparities (preferential treatment and reverse discrimination). The other two policies provided macro-level justifications in terms of achieving diversity or remedying historical injustices. We predicted that respondents would show less resistance to policy statements with explicit macro-level justifications than to policy statements that focus on the micro-level of implementation. In addition, to evaluate the possibility that resistance to affirmative action may be an expression of racial bias, we assessed participants' reactions to affirmative action policies involving three target groups: blacks, elderly persons, and handicapped persons. To the extent that racial bias is a key factor in reactions to affirmative action, white participants would be expected to exhibit more negative responses to policies targeted at blacks than at other groups. Moreover, based on the hypothesis that aversive racism motivates opposition to affirmative action, we predicted that

attitudes toward affirmative action would be particularly negative when blacks were the salient beneficiaries and white respondents could justify their opposition on the basis of some factor other than race, such as violations of procedural fairness.

Our results underscore the importance of the manner in which a policy is framed in shaping public opinion. Programs that were framed in terms of macro-justice by remedying historical injustice (past discrimination) or increasing cultural diversity were more acceptable to respondents than were those that focused on specific implementation (i.e., preferential treatment and reverse discrimination). In addition, consistent with the hypothesis that racism contributes to resisting affirmative action, policies directed at benefiting blacks yielded generally more negative responses than policies for persons with physical disabilities or elderly persons. Moreover, supportive of the specific predictions of the aversive racism framework, whites' responses to affirmative action were particularly negative when the group described as benefiting was blacks (vs. disabled or elderly persons) *and* the goal of the policy was presented as involving preferential treatment or reverse discrimination (vs. achieving diversity or compensating for past discrimination).

A second experiment (Dovidio & Gaertner, 1996) further examined the effects of framing and explicitly examined the mediating role of perceived fairness in responses to affirmative action benefiting different groups. In this experiment, affirmative action was framed at a macro level (i.e., to correct for past injustice) or at a micro level (i.e., positive action for the individual). We also manipulated the salience of the group associated with the policy: blacks, handicapped persons, or Native Americans. Native Americans were included because pretesting indicated that they represented a minority group that did not evoke significant negative reactions among potential participants.

Consistent with the results of Murrell et al. (1994), we found, in general, that attitudes toward the policy were more negative when it was not framed in terms of macro-justice than when it was, and that attitudes were more negative when the group benefiting was blacks rather than Native Americans or handicapped people. Furthermore, as expected, the condition in which the policy was not framed in terms of correcting for past discrimination and it was described as benefiting blacks produced a uniquely high level of resistance. In addition, these effects were mediated by perceptions of perceived fairness. Thus, it was perceptions of unfairness of a policy that benefited blacks that provided the rationale for opposing affirmative action.

These types of biases also extend to other formal forms of decision making, such as juridic decisions and personnel selection. Decision making about legal issues is considered in the next section.

B. LEGAL DECISIONS

Traditionally blacks and whites have not been treated equally under the law (Sidanius, Levin, & Pratto, 1998). Across time and locations in the United States, blacks have been more likely to be perceived by jurors as guilty (Fairchild & Cowan, 1997), more likely to be convicted of crimes, and, if convicted, sentenced to longer terms for similar crimes, particularly if the victim is white (see Robinson & Darley, 1995). Although some evidence indicates that disparities in judicial outcomes are declining over time (Sommers & Ellsworth, 2001), aversive racism appears to have a continuing, subtle influence.

For example, Johnson, Whitestone, Jackson, and Gatto (1995) explored how the introduction of an apparently non–race-related factor suggesting guilt can differentially affect juridic decisions in ways that discriminate against black defendants. In particular, in a laboratory simulation study, Johnson et al. (1995) examined the impact of the introduction of inadmissible evidence, which 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 race when all the evidence presented was admissible. However, consistent with the aversive racism framework, the presentation of inadmissible evidence increased judgment of guilt when the defendant was black but not when the defendant was white. Furthermore, suggesting the unconscious or unintentional nature of the bias, participants' self-reports indicated that they believed that the inadmissible evidence had less effect on their decisions when the defendant was black than when the defendant was white. Johnson et al. (1995) conclude that these results "are clearly consistent with the modern racism perspective, which suggests that discriminatory behavior will occur only when it can be justified on nonracial grounds" (p. 896).

Several other studies of legal decision making have yielded evidence consistent with the proposition that whites' biases against blacks will be more pronounced when they have an apparently non–race-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 suggests that racial bias may be involved in the allegations against a black defendant, whites no longer racially discriminate (Sommers & Ellsworth, 2000).

We similarly found that providing white jurors with other types of justifications could also lead to discriminatory outcomes in capital sentencing (Dovidio, Smith, Donnella, & Gaertner, 1997). In particular, although aversive racists did not generally discriminate on the basis of race in their recommendations for the death penalty in a capital case, when they were

asked their judgments independently, they did recommend the death penalty for a black defendant significantly more strongly than for a white defendant when they learned that a black juror (but not a white juror) would vote for the death penalty. When a black juror also advocates the death penalty, it is easier for aversive racists to rationalize their vote for capital sentencing as not reflecting racial bias.

Another study of simulated juridic decisions by Faranda and Gaertner (1979) demonstrated how traditional and aversive forms of racism can combine to shape perceptions of a defendant's guilt. Specifically, this study investigated the hypothesis that, whereas the racial biases of those who are likely to have traditionally racist attitudes (high authoritarians) would reflect primarily anti-black biases, the racial biases of those who are likely to exhibit aversive racism (low authoritarianism) would mainly represent pro-white biases. Thus, complementing the work of Johnson et al. (1995), this experiment examined the extent to which high- and low-authoritarian–scoring white college students playing the role of jurors would follow a judge's instruction to ignore inadmissible prosecution testimony that was damaging to a black or white defendant.

As predicted, both high- and low-authoritarian participants displayed racial biases in their reactions to the inadmissible evidence, but they did so in different ways. In their ratings of certainty of guilt, high authoritarians did not ignore the inadmissible testimony when the victim was black. They 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. That is, low authoritarians were less certain of the white defendant's guilt when the inadmissible evidence was presented than when it was omitted. Thus, lowauthoritarian participants demonstrated a pro–in-group bias. Importantly, the anti–out-group bias of high authoritarians and the pro–in-group bias of low authoritarians both disadvantage blacks relative to whites but in fundamentally different ways.

C. SELECTION DECISIONS

As noted at the beginning of this chapter, labor statistics continue to demonstrate fundamental disparities in the economic status of blacks relative to whites—a gap that has not only persisted but also, in some important

aspects such as family income, has widened in recent years (see Blank, 2001). Aversive racism may be one factor that contributes to disparities in the workplace by influencing both the access of blacks to the workplace and their performance in it.

At the time of hiring, aversive racism can affect how qualifications are perceived and weighed in a manner that systematically disadvantages black relative to white applicants. In particular, the aversive racism framework suggests that bias will not be expressed when a person is clearly qualified or unqualified for a position, because the appropriate decision is obvious. However, bias is expected when the appropriate decision is unclear, for example, when it is not clear whether the candidate's qualifications meet the criteria for selection or when the candidate's file has conflicting evidence (e.g., some strong and some weak aspects).

In one study of hiring decisions (Dovidio & Gaertner, 2000), we presented college students with excerpts from an interview and asked them to evaluate candidates for a position in an ostensibly new program for peer counseling at their university. Specifically, white participants evaluated a black or white candidate who had credentials that were systematically manipulated to represent very strong, moderate, or very weak qualifications for the position. These findings were supportive of 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), there was no discrimination 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 (see Table I).

In subsequent research (Hodson, Dovidio, & Gaertner, 2002), participants were asked to help make admissions decisions for the university. Given the social climate on college campuses today, it is possible that even higher prejudice-scoring students may be concerned about viewing themselves as prejudiced. Consequently, as we have observed among lower prejudiced participants in the past, these individuals may currently express their negative attitudes in subtle, indirect, and rationalizable ways—and, relative to the general population, these higher prejudice-scoring college students may actually be low to moderate in prejudice and not view themselves as racially prejudiced. Indeed, among a comparable sample of higher

TABLE I

RECOMMENDATIONS AND ATTRIBUTIONS OF PERSONAL CHARACTERISTICS AND INTERPERSONAL
ORIENTATION AS A FUNCTION OF CANDIDATE QUALIFICATIONS AND RACE

Condition	Strength of recommendation*			Percent recommended		
	1988–1989	1998–1999	Both	1988–1989	1998–1999	Both
Strong qualifications						
White candidate	6.74 (1.41)	6.21 (2.09)	6.52 (1.72)	89%	79%	85%
Black candidate	7.32 (1.67)	7.00 (1.60)	7.18 (1.62)	95%	87%	91%
Moderate qualifications						
White candidate	6.05 (1.73)	5.69 (1.60)	5.91 (1.67)	75%	77%	76%
Black candidate	5.06 (1.39)	4.53 (1.64)	4.82 (1.51)	50%	40%	45%
Weak qualifications						
White candidate	3.05 (1.65)	2.42 (1.68)	2.81 (1.66)	5%	8%	6%
Black candidate	3.29 (1.69)	3.77 (1.69)	3.50 (1.68)	12%	15%	13%

*Values are expressed as from Dovidio & Gaertner, 2000. Means with standard deviation presented in parentheses.

prejudice-scoring participants, only 15% regarded themselves as "prejudiced against blacks."

We found no anti-black bias among our higher and lower prejudice-scoring college participants when applicants had uniformly strong or uniformly weak college board scores and records of high school achievement. However, when applicants were strong on one dimension (e.g., on college board scores) and weak on the other (e.g., high school grades), black applicants tended to be recommended less strongly than were white applicants among higher prejudice scoring-participants. Moreover, these participants systematically changed how they weighed the criteria to justify their decisions as a function of race. For black applicants, higher prejudice-scoring college participants gave the weaker dimension (college board scores or grades) greater weight in their decisions, whereas for white applicants they assigned the stronger of the qualifications more weight. Analogously, Brief, Dietz, Cohen, Pugh, and Vaslow (2000) found that white interviewers who scored high on the Modern Racism Scale (McConahay, 1986) were particularly likely to discriminate against black relative to white applicants in hiring when a business-related justification for not hiring the candidate was available. Taken together, these findings suggest that when given latitude for interpretation, higher prejudice white college participants (whom, relative to the general population may be regarded as generally

moderate to low prejudiced, see Schuman et al., 1997), give white candidates the "benefit of the doubt," a benefit they do not extend to blacks.

D. SUBTLE BIAS: A SUMMARY

In summary, the behavior of aversive racists is characterized by two types of inconsistencies. First, aversive racists exhibit an apparent contradiction between their expressed egalitarian attitudes and their biased (albeit subtle) behaviors. Second, sometimes (in clear situations) they act in an unbiased fashion, whereas at other times (in ambiguous situations) they are biased unintentionally against blacks.

Overall, we have offered evidence across time, populations, and paradigms that illustrates how aversive racism—racism among people who are good and well intentioned—can produce disparate outcomes between blacks and whites. As noted earlier, although the bias of aversive racists may be subtle and unintentional, its consequences may ultimately be just as severe as old-fashioned racism—threats to the well-being of blacks and the restriction of opportunities.

Furthermore, the racial biases of aversive racists are often manifested in terms of a pro–in-group rather than an anti–out-group bias (Gaertner et al., 1997). A pro–in-group bias is often less readily recognized, and when it is this type of bias is often less threatening to one's nonprejudiced self-image than is overt bias against a black person. For instance, we found that white college students did not overtly associate negative characteristics more strongly with blacks than with whites, responses that might be interpreted as reflecting anti-black attitudes, but they did associate positive character-istics more strongly with whites than with blacks (Dovidio & Gaertner, 1991; Gaertner & McLaughlin, 1983). This is not the old-fashioned, overt type of bias associated with the belief about black inferiority but instead is a modern, subtle form of bias that reflects a belief about white superiority. In addition, pro–in-group bias is typically not encompassed in legal definitions of discrimination (Krieger, 1998).

However, the distinction between anti-black and pro-white responses can have important implications. First, it may provide a more comprehensive and accurate understanding of contemporary racism. The conclusions drawn from considering only the anti–out-group portion of these attitudes in isolation of pro–in-group attitudes might misrepresent the overall phe-nomenon. For example, Crocker and Schwartz (1985) found that when looking at only out-group attitudes, people with low self-esteem appeared more prejudiced than those with high self-esteem. However, when both in-group and out-group attitudes were considered, people with high

elf-esteem were more *biased* (i.e., evaluated the out-group more negatively han the in-group) than those with low self-esteem. Second, pro-white atti- udes may be one example of the broader class of phenomena, which ve considered earlier in our discussion of social categorization, in which people generally favor in-group members over out-group members. Con- ceiving of black–white relations within a more general context of intergroup relations (while still recognizing the unique cultural and historical character- stics of the conflict that have shaped sterotypes and status relations) has heoretical and practical advantages for identifying factors that perpetuate such biases and the factors that may increase intergroup harmony.

In general, then, we have obtained substantial evidence indicating the existence of aversive racism and demonstrating how it affects interracial behavior. However, because aversive racists are guarded about appearing prejudiced, to others and to one's self, they may consciously or unconscious- ly alter their responses to appear nonprejudiced, particularly in contexts in which race or racial attitudes are salient. As a consequence, aversive racists often appear nonprejudiced, in an absolute sense, on self-report measures of prejudice. Recent advances in attitude measurement, particularly in terms of the assessment of attitudes and beliefs that are out of conscious awareness (e.g., implicit attitudes) have permitted a closer examination of how the conscious and unconscious forces hypothesized within the aversive racism framework operate. We consider these developments in the next section.

IV. Dissociated Attitudes

Beginning with our earliest work on the aversive racism framework, we hypothesized that a dissociation commonly exists between whites' conscious and unconscious racial attitudes and beliefs. Recent research in social cog- nition has yielded new techniques for assessing unconscious, as well as conscious, attitudes and stereotypes. These techniques thus provide direct evidence about the influence of factors previously only assumed to be involved in aversive racism.

A. IMPLICIT PROCESSES

Borrowing from work in cognition more generally, researchers have made a fundamental distinction between explicit and implicit processes (Devine, 1989; Greenwald & Banaji, 1995). Explicit attitudes and stereotyping operate in a conscious mode and are exemplified by traditional, self-report measures of these constructs. In contrast, implicit attitudes and stereotypes 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 and unintentional 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 (see Blair, 2001; Dovidio et al., 2001).

We, along with other researchers using response-time measures based on the assumption that racial attitudes operate like other stimuli to facilitate responses and decision making about related concepts (e.g., doctor–nurse), have found consistent evidence of whites' generally negative implicit (unconscious) attitudes toward blacks (e.g., Dovidio et al., 1986; Fazio, Jackson, Dunton, & Williams, 1995; Gaertner & McLaughlin, 1983; Greenwald, McGhee, & Schwartz, 1998; Wittenbrink, Judd, & Park, 1997).

A study from Dovidio, Kawakami, Johnson, Johnson, and Howard (1997) illustrates a paradigm we have used to assess unconscious attitudes. In this technique, we do not mention race, which might prepare our participants to censor negative feelings. Instead, we presented black and white primes (schematic faces) subliminally. On each of the key experimental trials, we first presented a sketch of a black or white person very rapidly on a computer screen, and then we covered up ("masked") the sketch with the large letter "P" within an oval (to indicate a trial on which participants would make decisions about people) or "H" (to indicate a trial on which participants would make decisions, in control trials, about houses) in the same area of the computer screen so that participants are unaware that the sketch even appeared. Thus participants were not cognizant of the fact that we were assessing their racial beliefs and feelings. On the trials in which the letter "P" was clearly visible to participants, they were asked to make a decision about whether the next word that appears could ever describe a person (i.e., P for person). Next, we displayed a characteristic, for example "good" or "bad," and recorded how long participants took to make the decision. Faster response times are assumed to reflect greater association. We then examined whether the subliminal sketch of a black or white person would affect their decision-making times.

Using this subliminal procedure we found that our white participants had more positive associations with whites than with blacks and more negative associations with blacks than with whites, even though they were not aware of the schematic faces or that the study tested their racial attitudes. These findings converge with a substantial number of studies (see Blair, 2001; Dovidio et al., 2001) using a broad range of techniques (such as the Implicit Association Test; Greenwald et al., 1998) that reveal that the vast majority of white Americans harbor unconscious negative associations about blacks.

Moreover, supportive of the aversive racism framework, whites' uncon-
scious attitudes are largely dissociated from their conscious, self-reported
attitudes but somewhat less so when the motivation to respond in socially
desirable ways on self-reported measures of racial attitudes is reduced (Nier,
2003). The correlation between these different types of attitudes is, on
average, 0.24 (Dovidio et al., 2001). Nevertheless, the development of these
new techniques thus allows us to examine the independent influence of
conscious and unconscious attitudes to whites' behaviors toward blacks, as
well as their joint influence.

We hypothesize that this disassociation between the conscious (explicit)
and unconscious (implicit) attitudes of aversive racists can subtly shape the
ways that whites and blacks interact and further contribute to the different
perceptions that whites and blacks develop about their situations. If whites
are unaware of their negative implicit attitudes, they may also be unaware of
how their behaviors in interracial interactions may be influenced by these
racial biases. In contrast, blacks, who can observe the negative behaviors of
whites with whom they are interacting, may form very different impressions
about whether racial bias is operating and the degree to which it is inten-
tionally determined. Blacks (and other minority groups) may be vigilant to
signs of bias and readily attribute these actions to intentional racism
(Shelton, 2000; Vorauer & Kumhyr, 2001). We examine the implications of
this aspect of our framework in the next section.

B. CONFLICTING ATTITUDES, MIXED MESSAGES

We propose that the dissociation between the positive conscious attitudes
and the negative unconscious attitudes of aversive racists fundamentally
relates to the ways they interact with blacks. In particular, conscious
and unconscious attitudes influence behavior in different ways and under
different conditions (Dovidio, Kawakami, & Gaertner, 2002; Dovidio,
Kawakami, et al., 1997; Fazio & Olson, 2003; Fazio et al., 1995; Wilson,
Lindsey, & Schooler, 2000). Conscious 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. Unconscious
attitudes influence responses that are more difficult to monitor and control
(e.g., some nonverbal behaviors; see Chen & Bargh, 1997; McConnell &
Leibold, 2001) or responses that people do not view as an indication of their
attitude and thus do not try to control. For instance, we have found that whites'
unconscious negative attitudes predict nonverbal cues of discomfort (increased
rate of blinking) and aversion (decreased eye contact) toward blacks (see also
Word, Zanna, & Cooper, 1974), whereas whites self-reported, conscious

attitudes predict overt evaluations and indications of liking toward blacks (Dovidio, Kawakami et al., 1997). Thus, aversive racists, who have positive conscious attitudes and who want to be supportive of blacks but who also harbor unconscious negative attitudes or associations (see Karpinski & Hilton, 2000), are likely to convey mixed messages in interracial interactions. Given these conflicting signals, it is not surprising that blacks are likely to approach interracial interactions with anxiety, guardedness, and underlying mistrust (Hyers & Swim, 1998; Shelton, 2000).

C. INTERRACIAL INTERACTION AND IMPLICIT AND EXPLICIT ATTITUDES

These potential communication obstacles and interaction problems are exacerbated by the fact that whites and blacks have fundamentally different perspectives on the attitudes implied and the actions demonstrated by whites during these interactions. Whites have full access to their conscious attitudes and are able to monitor and control their more overt and deliberative behaviors. They do not have access to their unconscious attitudes or to their less monitorable behaviors. As a consequence, whites' beliefs about how they are behaving or how blacks perceive them would be expected to be based primarily on their conscious attitudes and their more overt behaviors, such as the verbal content of their interaction with blacks, and not on their unconscious attitudes or less deliberative (i.e., nonverbal) behaviors. In contrast to the perspective of whites, the perspective of black partners in these interracial interactions allows them to attend to both the spontaneous (e.g., nonverbal) and the 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 than are whites (Shelton, 2000).

To investigate this possibility, we conducted another experiment (Dovidio et al., 2002). We assessed perceptions of interracial interactions by whites and blacks, and we related those perceptions to white participants' conscious attitudes, measured on a self-report prejudice scale, and unconscious attitudes, assessed with a response-latency technique. Then we arranged interracial conversations with a black and a white dyad partner around a race-neutral topic. We videotaped the interactions and subsequently had one set of coders rate the nonverbal and verbal behaviors of white participants and another set of observers rate their global impressions of participants from a videotape recorded from their partners' perspective.
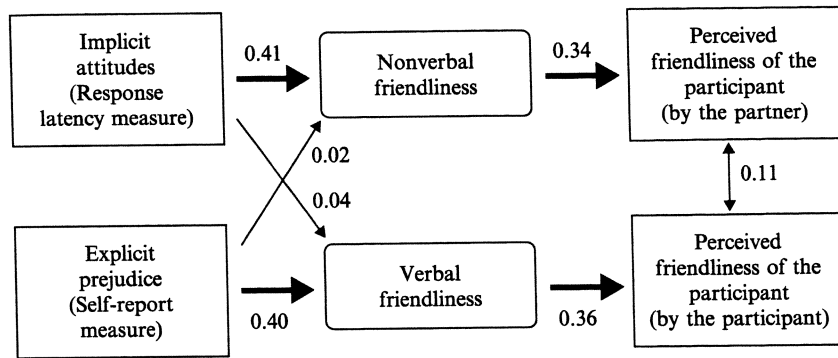
Fig. 1. The relationships (correlations) between measures of prejudice and participant behavior and impressions. From Dovidio, Kawakami, & Gaertner, 2002.

As we hypothesized (see Fig. 1), white participants' self-reported racial attitudes predicted their deliberative behaviors such as their verbal friendliness toward black relative to white partners, which in turn predicted white participants' impressions of how friendly they behaved in interactions with the black relative to the white partner. Thus, whites participants' conscious attitudes, controllable behaviors, and self-impressions were all consonant. Unconscious racial attitudes, measured with response latencies, did not predict these verbal behaviors or white participants' impressions of how they behaved. However, as we also anticipated, we found that white participants' unconscious racial attitudes reflected by their response latencies predicted biases in their *nonverbal* behaviors (as scored by our observers), which then predicted how they were perceived by their partners (see Fig. 1).

Because white participants and their partners based their impressions on different aspects of the participants' attitudes, the conscious and unconscious attitudes were dissociated and their impressions of the interaction were generally uncorrelated ($r = 0.11$). White participants typically reported that they found the interaction satisfying, and they expressed contentment with their contributions. Their black partners, however, reported being relatively dissatisfied with the exchange and were uneasy about their partners' behaviors. Despite white participants' good intentions, the impressions they made were not as good as they thought. Moreover, both dyad members assumed that their partner shared the impression of the interaction as they did.

Taken together, our findings on the effects of conscious and unconscious attitudes in interracial interaction suggest that the nature of contemporary biases can shape the everyday perceptions of white and black Americans in ways that interfere with the communication and trust that are critical to

developing long-term positive intergroup relations. These different perspectives and experiences of whites and blacks in interracial interaction, which happen inadvertently and occur daily, can have summative effects over time (Feagin & Sikes, 1994) and help to contribute to the climate of misperception and distrust that characterizes contemporary race relations in the United States. The majority of blacks in America today have a profound distrust for the police and legal system, and about a third are overtly distrustful of whites in general (Anderson, 1996). In addition, blacks commonly believe that conspiracies inhibit the progress of blacks (Crocker, Luhtanen, Broadnax, & Blaine, 1999). The mixed messages that aversive racists often convey can create fundamental miscommunication in interracial interaction and produce divergent impressions among interactants that can undermine their ability to interact efficiently in task-oriented situations as well as effectively in social situations.

D. INTERRACIAL PERFORMANCE

The different and potentially divergent impressions formed by blacks and whites during interracial interactions can have significant impact on their effectiveness in task-oriented situations. Cannon-Bowers and Salas (1999) have argued that effective teamwork requires two types of skills, those associated with the technical aspects of the job and those associated with being a member of the team. For this latter factor, team competencies include the knowledge, skills, and attitudes required to work effectively with others. In addition to manifesting itself in terms of different impressions and perceptions, contemporary bias can influence personal relations and group processes in ways that unintentionally but adversely affect outcomes for blacks.

We examined these processes in interracial dyads in which a black participant was paired with a white student who was identified as a traditionally high-prejudiced person (who expressed their bias openly), an aversive racist (who expressed egalitarian views but who showed evidence of unconscious bias), or a low-prejudiced white (who held egalitarian views and showed little evidence of unconscious bias) (Dovidio, Gaertner, Kawakami, & Hodson, 2002). These participants engaged in a problem-solving task about challenges to college students. For example, in one task they were asked to identify the five most important things that incoming students need to bring to campus. Because there were no objective measures of the quality of their team solution, we focused on the quality of their interaction (as reflected in their perceptions of friendliness and trustworthiness and feelings of satisfaction) and on their efficiency (as indexed by their time to complete the task).

In general, whites' impressions of their behavior were related primarily to their self-reported expressed attitudes, whereas blacks' impressions of whites were related mainly to whites' unconscious attitudes. Specifically, whites who expressed egalitarian ideals (i.e., low-prejudiced whites and aversive racists) reported that they behaved in more friendly ways than did those who expressed their bias openly (i.e., high-prejudiced whites). Black partners perceived only whites who showed no evidence of unconscious bias (i.e., low-prejudiced whites) to be more friendly than those who had unconscious biases (aversive racists and high-prejudiced whites). Of all three groups, blacks were least trustful of aversive racists.

Our results further revealed that whites' racial attitudes could be systematically related to the efficiency of the interracial teams. Teams with low-prejudiced whites solved the problem most quickly. Interracial teams involving high-prejudiced whites were next most efficient. Teams with aversive racists were the least efficient. Presumably, the conflicting messages displayed by aversive racists and the divergent impressions of the team members' interaction interfered with the team's effectiveness. To the extent that blacks are in the minority in an organization and are dependent on high-prejudiced whites or aversive racists, their performance is likely to be objectively poorer than the performance of whites who predominantly interact with other whites. Thus, even when whites harbor unconscious and unintentional biases toward blacks, their actions can have effects sometimes even more detrimental than those of overt racists on interracial processes and outcomes.

Overall, we have offered a range of evidence across time, populations, and paradigms that illustrates how aversive racism—racism among people who are good and well intentioned—can influence the nature of interracial interactions and directly or indirectly produce disparate outcomes between blacks and whites. As noted earlier, although the bias of aversive racists may be subtle and unintentional, its consequences may ultimately be just as debilitating, for example, by creating barriers to their advancement in employment settings, to blacks as old-fashioned racism. In the next section we examine strategies for combating this insidious type of bias.

V. Combating Aversive Racism

When we describe our findings formally, in papers and presentations, and informally, a question often arises, "What can we do about subtle biases, particularly when we do not know for sure whether we have them?" Like a

mutating virus, racism may have evolved into different forms that are more difficult not only to recognize but also to combat.

Traditional prejudice-reduction techniques have been concerned with changing conscious attitudes—old-fashioned racism—and obvious expressions of bias. Attempts to reduce this direct, traditional form of racial prejudice have typically involved educational strategies to enhance knowledge and appreciation of other groups (e.g., multicultural education programs), emphasize norms that prejudice is wrong, and involve direct (e.g., mass media appeals) or indirect (dissonance reduction) attitude change techniques (Stephan & Stephan, 2001). However, because of its pervasiveness, subtlety, and complexity, the traditional techniques for eliminating bias that emphasized the immorality of prejudice and illegality of discrimination are not effective for combating aversive racism. Aversive racists recognize that prejudice is bad, but they do not recognize that *they* are prejudiced.

We believe, however, that aversive racism can be addressed with techniques aimed at its roots at both individual and collective levels. At the individual level, strategies to combat aversive racism need to be directed at unconscious attitudes. Aversive racists' conscious attitudes are already favorable, and may, in fact, be instrumental in motivating change. At the intergroup level, interventions may be targeted at processes that support aversive racism, such as ingroup favoritism (Gaertner & Dovidio, 2000).

A. ADDRESSING UNCONSCIOUS ATTITUDES AND BELIEFS

Aversive racism is characterized by conscious (explicit) egalitarian attitudes and negative unconscious (implicit) attitudes and beliefs. Wilson et al. (2000) argue that systems of dual attitudes, such as those involved in aversive racism, typically arise developmentally. The person's original attitudes, through repeated occurrence, practice, and ultimately overlearning, become unconscious and automatically activated (Wyer & Hamilton, 1998). Given the historic socialization of whites and the repeated exposure to negative images of blacks in the mass media, most whites develop negative attitudes and stereotypical belief about blacks that become internalized and habitualized relatively early in life (Devine, 1989). Aversive racists, however, also subsequently develop a strong conscious commitment to equality and to being nonprejudiced. Nevertheless, according to Wilson et al.'s (2000) model, the original attitude is not replaced. It is stored in memory and is implicit and unconscious, while the newer attitude is explicit and conscious. In general, explicit attitudes can change and evolve relatively easily, whereas implicit attitudes, because they are based in overlearning and habitual reactions, are much more difficult to alter.

Just because they are unconscious and automatically activated does not mean that aversive racists' unconscious negative attitudes are immutable and inevitable. If unconscious attitudes and stereotypes can be learned, we propose that they can also be unlearned or inhibited by equally well-learned countervailing influences. Devine and Monteith (1993) observed, "Although it is not easy and clearly requires effort, time, and practice, prejudice appears to be a habit that can be broken" (p. 336). We have found that with extensive practice, either imposed externally or self-motivated, it is possible to change implicit beliefs.

1. Imposed Practice

With extensive practice, individuals can develop "auto-motive" control of their actions through frequent and persistent pursuit of a goal, such as to not be biased or not to stereotype (Bargh, 1990). As Monteith, Sherman, and Devine (1998) note, "Practice makes perfect. Like any other mental process, thought suppression processes may be proceduralized and become relatively automatic" (p. 71).

Consistent with this line of reasoning, we found in a series of studies (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000) that automatic stereotype activation can be reduced and eliminated with training to not stereotype members of a group. In particular, participants in this research practiced extensively to respond in ways either consistent with prevailing racial stereotypes (by indicating "yes" to stereotype-consistent pairings of black and white photographs and traits and responding "no" to stereotype-inconsistent pairings) or to negate racial stereotypes (by responding "no" to stereotype-consistent pairings and "yes" to stereotype-inconsistent pairings). At the end of the session, participants performed a response-latency task to assess their unconscious, automatically activated racial stereotypes. Whereas those participants who were in the condition in which they responded affirmatively to conventional stereotypic associations showed equivalent evidence of unconscious racial stereotypes before and after the training exercise, those who practiced negating stereotypes demonstrated a significant decrease in unconscious stereotyping after training. These effects of practice in negating stereotypes were also still evident for 24 hours after the training.

Although such direct strategies appear to be promising, these kinds of intensive and time-consuming approaches may be limited in their general applicability. Alternative promising strategies, however, take advantage of aversive racists' genuine interest in being nonprejudiced to motivate significant and enduring change.

2. Motivation and Self-Regulation

Because aversive racists consciously endorse egalitarian values and truly want to be nonprejudiced, it may be possible to capitalize on their good intentions and induce self-motivated efforts to reduce unconscious biases on becoming aware of them. Work by Monteith, Devine, and their colleagues (e.g., Devine & Monteith, 1993; Monteith & Voils, 1998) has revealed that when low-prejudiced people recognize discrepancies between their potential behaviour toward minorities (i.e., what they *would* do) and their personal standards (i.e., what they *should* do), they feel guilt and compunction, which produces motivations to respond without prejudice in the future. In their process model of prejudice reduction, Devine and Monteith (1993) further suggest that individuals who are committed to maintaining egalitarian standards learn to reject old, biased ways of responding and to adopt new, nonprejudiced ways. Over time and with practice, these people learn to reduce prejudicial responses and to respond in ways that are consistent with their nonprejudiced personal standards. Thus, this process of self-regulation, which is initiated by making people aware of their potential for racial bias, may produce changes in even unconscious negative responses when extended over time.

We directly investigated this possibility (see Dovidio, Kawakami, & Gaertner, 2000). White participants, who were categorized as low or high in prejudice on the basis of their self-reported prejudice, completed a task making them aware of discrepancies between what they would do and what they should do (i.e., their personal standards) in interracial situations (Devine & Monteith, 1993). We assessed emotional reactions and, using a response-latency task, initial unconscious racial stereotyping. Three weeks later participants returned to the laboratory and completed the unconscious stereotyping tasks and another measure of the "would–should" discrepancy.

We hypothesized that initial discrepancies between one's actions (what one would do) and personal standards (what one should do) would generate stronger feelings of guilt and compunction and produce more self-initiated efforts at change among low-prejudiced than among high-prejudiced participants. The effects of this self-regulatory process were expected to be reflected in decreased discrepancies and unconscious stereotyping.

As anticipated, greater discrepancies between what one would do and should do produced higher levels of guilt in the first session, and this relationship occurred primarily for low-prejudiced participants. These findings indicate the potential initiation of self-regulatory processes for low- but not high-prejudiced participants. When participants returned 3 weeks later, we found an overall greater alignment (i.e., smaller discrepancy) between what one would and should do—an indication that both high- and

low-prejudiced participants showed a decrease in overt expressions of bias. However, as hypothesized, low- and high-prejudiced whites differed in terms of the extent to which they internalized these changes. Low-prejudiced whites who had larger initial discrepancies showed greater reductions in unconscious stereotyping ($r = -0.56$); in contrast, for high-prejudiced whites the relationship was weaker (-0.07) and nonsignificant. These findings demonstrate that the good intentions of aversive racists can be harnessed to produce self-initiated change in even unconscious biases with appropriate awareness, effort, and practice over time.

Recently Son Hing, Li, and Zanna (2002) have extended work along these lines by examining responses to being made aware of hypocrisy on participants' subsequent interracial responses. This study was conducted in Canada with Asians as the target group. Participants were classified on the basis of the combination of their responses to an explicit, self-report measure of prejudice toward Asians and a measure of implicit associations (spontaneous completion of words as Asian stereotypes). Participants who were low in explicit prejudice but who showed implicit biases were identified as aversive racists; those low in both explicit and implicit bias were considered truly low prejudiced.

Participants next wrote an essay, reinforcing their overt egalitarian orientation, about the importance of treating Asians fairly. Participants in the hypocrisy condition, which was designed to sensitize people to violations of their egalitarian principles, were asked to write briefly about two situations in which they reacted negatively or unfairly to an Asian person. Participants in a control condition were not asked to write about these situations. Participants then completed a mood questionnaire and, ostensibly after the study was completed, were asked to complete a survey for the university's student government on how funding cuts should be allocated to various campus groups. They were informed that a 20% cut in funding was needed for the budget of 10 campus groups. The main dependent measure was the amount of funding recommended for the Asian Students' Association.

In general, the results supported the prediction that making people aware of violations of their egalitarian principles would primarily arouse guilt among aversive racists who actually harbor negative feelings toward Asians, and thus produce compensatory behavior in recommended funding among aversive racists but not among nonprejudiced participants. Aversive racists in the hypocrisy condition experienced uniquely high levels of guilt and displayed the most generous funding recommendations for the Asian Students' Association. The funding recommendations of truly low-prejudiced participants were not affected by the hypocrisy manipulation. Son Hing et al. (2002) concluded that making people aware of their biases is particularly effective at reducing bias among people who explicitly endorse egalitarian

principles while also possessing implicit biases—the factors that characterize aversive racists.

Strategies that emphasize intergroup processes, such as intergroup contact and social categorization, represent alternative, complementary approaches to these individual-level approaches. We examine one such approach in the next section.

B. REDIRECTING IN-GROUP BIAS

One basic argument we have made in our research on aversive racism is that the negative feelings that develop toward other groups may be rooted, in part, in fundamental, normal psychological processes. One such process, identified in the classic work of Tajfel, Allport, and others, is the categorization of people into in-groups and out-groups, "we's" and "they's." As we noted earlier, social categorization, particularly in terms of in-groups ("we's) and out-groups ("they's"), is a fundamental process that contributes to aversive racism (Gaertner et al., 1997). In general, the mere categorization of people into ingroups and outgroups has a profound influence on social perception, affect, cognition, and behavior. Because race is a fundamental type of social categorization in the United States, race is associated with strong ingroup biases.

If bias is linked to fundamental, normal psychological processes, then attempts to ameliorate bias should be directed not at eliminating the process but rather at redirecting the forces to produce more harmonious intergroup relations. The process of social categorization is not completely unalterable. By shifting the basis of categorization from race to an alternative dimension we can potentially alter who is a *we* and who is a *they*, undermining a contributing force to aversive racism.

Categories are hierarchically organized, with higher-level categories (e.g., nations) being more inclusive of lower-level ones (e.g., cities or towns). By modifying a perceiver's goals, motives, perceptions of past experiences, expectations, as well as factors within the perceptual field and the situational context more broadly, there is opportunity to alter the level of category inclusiveness that will be most influential in a given situation. This malleability of the level at which impressions are formed is important because of its implications for altering the way people think about members of ingroups and outgroups, and consequently about the ways whites in general, and aversive racists in particular, respond to blacks.

Because categorization is a basic process that is fundamental to intergroup bias, we have targeted this process as way of addressing the effects of aversive racism. The next section explores how the forces of categorization

can be harnessed and redirected toward the reduction, if not the elimination, of racial bias. This approach is represented by the Common Ingroup Identity Model (Gaertner & Dovidio, 2000; Gaertner, Dovidio, Anastasio, Bachman, & Rust, 1993).

1. The Common In-Group Identity Model

The Common In-Group Identity Model is rooted in the social categorization perspective of intergroup behavior and recognizes the central role of social categorization in both reducing and creating intergroup bias (Tajfel & Turner, 1979). Specifically, if members of different groups are induced to conceive of themselves more as a single, superordinate group rather than as two separate groups, attitudes toward former out-group members will become more positive through processes involving pro–in-group bias. Thus, changing the basis of categorization from race to an alternative dimension can alter who *we* is and who *they* are, undermining a contributing force to contemporary forms of racism, such as aversive racism. Formation of a common identity, however, does not necessarily require groups to forsake their ethnic or other subgroup identities. It is possible for members to conceive of themselves as holding a "dual identity" in which both subgroup and superordinate groups are salient simultaneously. Substantial evidence across a variety of settings in support of the Common Ingroup Identity Model has been found (Gaertner & Dovidio, 2000).

In one test of the model, which investigated the causal role of common group identity in reducing bias (Gaertner, Mann, Murrell & Dovidio, 1989), members of two separate laboratory-formed groups were induced through various structural interventions (e.g., seating arrangement) either to maintain their original group identities (i.e., conceive of themselves as different groups) or to recategorize themselves as one superordinate group. As predicted, the manipulation to encourage recategorization of former out-group members within a common group identity produced more inclusive representations that ultimately mediated lower levels of intergroup bias, primarily by increasing the attractiveness of the former out-group members.

Additional research also shows that interventions that have been demonstrated to reduce prejudice, such as cooperative interaction and appropriately structured intergroup contact, reduce bias, at least in part, by altering the intergroup cognitive representations. With respect to cooperative interaction, we brought two three-person laboratory groups, which had worked separately on a problem-solving task, together under conditions designed to vary independently the members' representations of the six-person aggregate as one group or two groups and the presence or absence of intergroup cooperative interaction (Gaertner, Mann, Dovidio, Murrell, & Pomare, 1990).
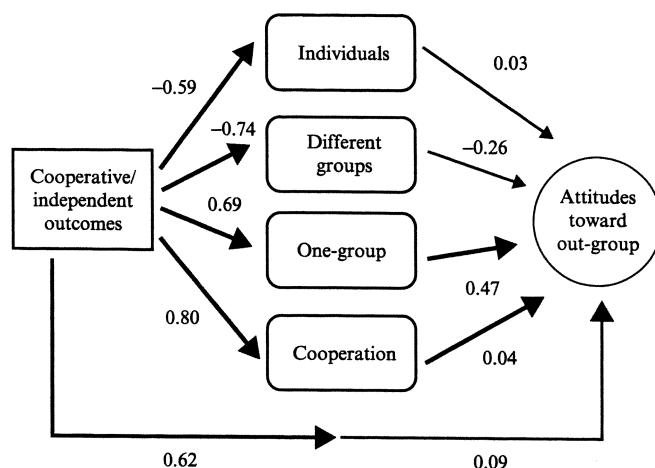
Fig. 2. The effect of cooperative interaction on improving attitudes toward an out-group occurs primarily by creating more inclusive, one-group representations Thick arrows indicate statistically significant paths ($P < 0.05$). From Gaertner et al., 1990.

As in our earlier experiment (Gaertner et al., 1989), our manipulation designed to influence representations of the aggregate as one group did produce stronger one-group representations, which in turn mediated lower degrees of intergroup bias in evaluations of original in-group and out-group members. Also, consistent with our hypothesis, when separate group identities were initially salient, cooperative interaction (see Fig. 2) decreased the extent to which the aggregate felt like two groups and increased the extent it felt like one. The inclusive, one-group representation then led to more favorable evaluations of out-group members, which contributed to reduced intergroup bias.

Using survey techniques under more naturalistic circumstances, we also found converging support for the proposition that the features specified by the Contact Hypothesis (Allport, 1954; Williams, 1947) reduce intergroup bias, in part, because they transform members' representations of the memberships from separate groups to a single, more inclusive group. Participants in these studies included students attending a multiethnic high school (Gaertner, Rust, Dovidio, Bachman & Anastasio, 1996), banking executives who had experienced a corporate merger involving a wide variety of banks across the United States (Bachman, 1993), and college students from blended families whose households are composed of two formerly separate families trying to unite into one (Banker & Gaertner, 1998).

Consistent with the role of an inclusive group representation that is hypothesized in the Common In-group Identity Model, across all three

studies (1) conditions of intergroup contact that were perceived as more favorable predicted lower levels of intergroup bias, (2) more favorable conditions of contact predicted more inclusive (one group) and less exclusive (different groups) representations, and (3) more inclusive representations mediated lower levels of intergroup bias and conflict (see Gaertner, Dovidio, Nier, Ward, & Banker, 1999). Recently a longitudinal study of stepfamilies found evidence supportive of the direction of causality between the constructs proposed by our model across time (Banker, 2002). Thus, across a variety of intergroup settings and methodological approaches we have found basically strong and consistent support for the Common Ingroup Identity Model.

2. Reducing Racial Biases: Experimental Evidence

We have applied the general principles of the Common In-Group Identity to reducing racial biases in laboratory and field settings. Two studies reported by Nier, Gaertner, Dovidio, Banker, and Ward (2001) illustrate the effectiveness of this approach for addressing whites' biases toward blacks specifically. Another study (Houlette et al., 2004) explored a range of biases, including racial bias, among elementary school children.

In a laboratory experiment (Nier et al., Study 1), white college students participated in a session with a black or white confederate. These students were induced to perceive of themselves as separate individuals participating in the study at the same time or as members of the same laboratory team. The participants evaluated their black partners significantly more favorably when they were teammates than when they were just individuals without common group connections. In contrast, the evaluations of the white partner were virtually equivalent in the team and individual conditions. Thus, inducing a common in-group identity was particularly effective at producing positive responses toward blacks.

The second study (Nier et al., Study 2) was a field experiment conducted at the University of Delaware football stadium before a game between the University of Delaware and Westchester State University. Black and white students approached fans from both universities just before the fans entered the stadium. These fans were asked if they would be willing to be interviewed about their food preferences. Our student interviewers wore either a University of Delaware or Westchester State University hat. By selecting white fans wearing clothing that identified their university affiliation, we systematically varied whether fans and our interviewers had a common or different university identities in a context in which we expected university identities to be particularly salient. We predicted that making a common identity salient would increase compliance with the interviewer's request, particularly when the interviewer was black.

Supportive of predictions from the Common In-Group Identity Model, white fans were significantly more cooperative with a black interviewer when they shared a superordinate university identity than when they did not (60% vs. 38%). For white interviewers, with whom they already shared racial group membership, the effect was much less pronounced (43% vs. 40%). Thus, in field and laboratory settings, racial out-group members were accorded especially positive reactions when they shared common in-group identity with white participants relative to when the context did not emphasize their common group membership. These studies suggest the value of combating aversive racism at its roots, by strategically controlling the forces of ingroup favoritism that can produce subtle racial biases associated with aversive racism.

In a recent study (Houlette et al., 2004), we attempted to evaluate these principles further in the context of the Green Circle intervention program, which is designed to combat a range of biases (based on weight and sex, as well as race and ethnicity) with young children. The guiding assumption of the Green Circle Program, which is practically and theoretically compatible with the Common In-Group Identity Model, is that helping children bring people from different groups conceptually into their own circle of caring and sharing fosters appreciation of their common humanity, as well as respect for their differences. In particular, facilitators engage children in a variety of exercises designed to expand the circle. The facilitator points out that, "All of us belong to one family—the human family." Paralleling the Common In-Group Identity Model, Green Circle assumes that an appreciation of common humanity will increase children's positive attitudes toward people who would otherwise remain outside their circle of inclusion. First- and second-graders either participated in the Green Circle Program or were in a control group of classes that did not yet have the program.

In terms of outcomes, the Green Circle intervention motivated the children to be more inclusive with their most preferred playmate. Specifically, compared to children in the control condition who did not participate in Green Circle activities, those who were part of Green Circle showed significantly greater change in willingness to select other children who were different than themselves in race and in sex as a child that they "would most want to play with." These changes in the most preferred playmate involve a child's greater willingness to cross group boundaries in making friends—a factor that is one of the most potent influences in producing more positive attitudes toward the out-group as a whole (Pettigrew, 1998). In addition, these intergroup friendships can have cascading effects by reducing bias among peers. Making people aware that their friends have friends from another group also reduces prejudice toward the group as a whole (Wright, Aron, McLaughlin-Volpe & Ropp, 1997).

In summary, the experiments reviewed in this section show that creating a common group identity can combat a range of overt expressions of racial bias. In addition, other research has demonstrated that emphasizing common group membership can address other types of biases that are associated with aversive racism, such as orientations toward policies designed to benefit blacks and other traditionally disadvantaged groups. In a survey study of white adults, Smith and Tyler (1996, Study 1) measured the strength of respondents' superordinate identity as "American" and also the strength of their identification as "white." Regardless of whether they strongly identified with being white, those respondents with a strong American identity were more likely to base their support of affirmative action policies that would benefit blacks and other minorities on concerns about fairness for different groups rather than on self-interest or white group-interest. However, for those who identified themselves more strongly with being white than with being American, their position on affirmative action was determined more strongly by concerns regarding the *personal* impact of these policies. This pattern of findings suggests that a strong superordinate identity (such as being American) allows individuals to support policies that would benefit members of other racial subgroups without giving primary consideration to their own instrumental needs (see also Huo, Smith, Tyler, & Lind, 1996).

The next section considers how creating a common in-group identity can influence the basic motivational orientations and cognitive processes that form the basis for racial biases.

3. A Common In-Group Identity and the Motivational Orientations

Within the aversive racism framework, we propose that the negative feelings, beliefs, and behaviors often are expressed subtly and indirectly— in ways that are not readily attributable (by others or themselves) to racial bias and thus do not threaten an aversive racist's nonprejudiced self-image. From this perspective, a major motive of whites in interracial situations is to *avoid wrongdoing*. As we have demonstrated across a range of experiments, whites appear to monitor their interracial behaviors closely to avoid discriminating against blacks when norms for appropriate behaviors are clearly defined.

Moreover, these attempts by aversive racists to avoid wrongdoing appear to involve significant conscious effort. Richeson and Shelton (2003) found that whites high in implicit prejudice toward blacks performed more poorly on a cognitively demanding task after interacting with a black person than did whites low in implicit prejudice. Richeson and Shelton proposed that the cognitive effort required by high implicitly prejudiced whites to monitor their interracial behavior depleted their cognitive resources, resulting in a

decrement in performance on the subsequent task. Thus, whites can, at least under some circumstances and with some effort, successfully suppress negative beliefs, feelings, and behavior toward blacks when it is obvious that such expressions reflect racial bias.

Although motivations to avoid thinking, feeling, or behaving in a prejudicial way can have positive interracial consequences, such as limiting social conflict, it can also have unintended negative consequences. Besides depleting aversive racists' cognitive resources, efforts to avoid wrongdoing and suppress prejudicial thoughts have two important potential costs for interracial interactions. First, a concern about avoiding wrongdoing increases anxiety that can motivate avoidance or premature withdrawal from the interaction (as evidenced by Liberal Party members hanging up prematurely more frequently on black than on white callers, as described earlier; see Gaertner, 1973). This avoidant reaction precludes the opportunity for meaningful, self-revealing exchanges between in-group and out-group members (Hyers & Swim, 1998). Second, in view of recent work on stereotype suppression and rebound (e.g., Bodenhausen & Macrae, 1996), it is possible that once this self-imposed suppression is relaxed, negative beliefs, feelings, and behaviors would be even *more likely* to occur than if they were not suppressed initially.

In the search for strategies that could eliminate the indirect, rationalizable ways that aversive racists discriminate we have also considered the importance of establishing positive interpersonal and intergroup motivations rather than simply suppressing negative motivations. The Common In-group Identity Model, because it focuses on redirecting the forces of in-group favoritism, offers such promise. Specifically, the recognition of a common ingroup identity potentially changes the motivational orientation or intentions of aversive racists from trying to avoid wrongdoing to trying to *do what is right*.

Although this shift in orientation is subtle, it can have fundamental benefits. For instance, it may relieve intergroup anxiety (see Stephan & Stephan, 1985) and reduce the likelihood of negative consequences of effortful attempts to avoid wrongdoing, such as the increased accessibility of negative thoughts, feelings, and behavior that occur when suppression is relaxed (Monteith et al., 1998; Wegner, 1994). Some preliminary evidence from our laboratory suggests the potential promise of a common in-group identity to alter motivation in just such a positive way (Dovidio, Gaertner, & Kawakami, 1998; Gaertner & Dovidio, 2000).

In this experiment, white participants who were about to interact with a white or a black confederate were either asked to try to avoid wrongdoing, instructed to try to behave correctly toward the other person, informed that they were part of the same team with their partner and competing against a

team at a rival institution, or were given no instructions. The dependent measure of interest was the relative accessibility of negative thoughts, as assessed by changes in responses on a Stroop color-naming task after the interaction relative to responses on a baseline Stroop task administered before the interaction (see Lane & Wegner, 1995). A rebound effect would be reflected in greater accessibility (operationalized in terms of longer color-naming latencies) of negative relative to positive words on the posttest Stroop task.

We hypothesized that, because the primary motivation of aversive racists in interracial interaction is to avoid wrongdoing and thus to suppress negative thoughts and feelings, participants explicitly instructed to avoid wrongdoing and those given no instructions would show relatively strong accessibility of negative thoughts after interacting with a black confederate. In contrast, we expected participants instructed to behave correctly and those in the "same team" condition (who were hypothesized to adopt a positive orientation on their own) would escape such a rebound effect.

The results, while preliminary, are very encouraging. When the confederate was white, the experimental conditions did not differ significantly in the accessibility of negative thoughts from one another or from baseline. When the confederate was black, however, the increased accessibility of negative relative to positive characteristics (from the pretest to the posttest) in the avoid wrongdoing and no instructions conditions was significantly greater than in the do right and same team conditions, in which there was an increase in the accessibility of positive relative to negative thoughts. The pattern of these findings suggests that the development of a common in-group identity can alter motivation in interracial situations from one of suppressing negative thoughts, feelings, and actions to one that is positive, more appetitive, and prosocial—and in a way that does not ironically result in further increases in negative thoughts. These findings are particularly encouraging to us because they illustrate the effectiveness of the Common In-group Identity Model for addressing individual-level biases and particularly the underlying dynamics of aversive racism.

VI. Summary and Conclusions

This chapter has described the concept of aversive racism, considered the factors contributing to aversive racism, demonstrated empirically how it affects outcomes for blacks and shapes interracial interactions, and explored how it can be combated. Despite apparent consistent improvements in expressed racial attitudes over time, aversive racism continues to exert a

subtle but pervasive influence on the lives of black Americans. This bias is expressed in indirect and rationalizable ways that restrict opportunities for blacks while insulating aversive racists from ever having to confront their prejudices.

Most of the work reviewed in this chapter has focused on the influence of contemporary racial biases of whites toward blacks because of the central role that racial politics has played in the history of the United States. Within the United States, this dynamic also relates to varying degrees to orientations toward women (Dovidio & Gaertner, 1983b; Rudman & Kilianski, 2000) and more recently to homosexuals (Hebl, Foster, Mannix, & Dovidio, 2002). Moreover, we propose that many of the principles of aversive racism also apply more generally to the responses of the majority group to minority groups in contexts in which egalitarian ideals are valued and discrimination is censured (Dovidio, Gaertner, Anastasio, & Sanitioso, 1992). For instance, Pettigrew and Meertens (1995) have found that whereas blatant prejudice in Europe is related to the unconditional exclusion or severe limitation of immigrants, subtle prejudice is associated with constraints in immigration, such as prerequisite educational levels, that can be justified on the basis of factors ostensibly unrelated to race or ethnicity. However, paralleling the scenario that we described for race relations, when conditions change and people feel threatened—such as after the terrorist attacks of September 11, 2001, in the United States—subtle biases may become more open and result not only in incidents of overt violence but also directly in attitudes and political actions against immigrants and immigration (Esses, Dovidio, & Hodson, 2002).

Bias associated with aversive racism is an elusive phenomenon, and the situation plays a critical moderating role. When an interracial situation is one in which an action could be readily attributed to racial bias, aversive racists carefully monitor their interracial behaviors and do not discriminate. In fact, they may respond even more favorably to blacks than to whites as a way of affirming their nonprejudiced self-images. But when the situation is ambiguous, norms for appropriate behavior are not clear, the circumstances permit a justification for negative behavior on the basis of some factor other than race, or aversive racists are not conscious of their actions, their bias is expressed.

The challenge for addressing aversive racism thus resides in its elusiveness. Because aversive racists are unaware of their unconscious negative attitudes and their effects, and truly embrace their egalitarian self-image, they are motivated to deny the existence of these feelings and not to recognize or take responsibility for the adverse impact of their behavior on blacks. The subtle processes underlying discrimination motivated by aversive racism can be identified and isolated under the controlled conditions of the laboratory;

however, at the societal and organizational levels, at which the controlled conditions of an experiment are rarely possible and multiple factors may shape decision making simultaneously, this process presents a substantial challenge to the equitable treatment of members of disadvantaged groups. Krieger (1995), in the *Stanford Law Review*, observed: "Herein lies the practical problem. ... Validating subjective decisionmaking systems is neither empirically nor economically feasible, especially for jobs where intangible qualities, such as interpersonal skills, creativity, and ability to make sound judgments under conditions of uncertainty are critical" (p. 1232). Thus the operation of aversive racism may go largely unnoticed and unaddressed in naturalistic settings.

In addition, to the extent that discrimination reflects in-group favoritism (see also Gaertner et al., 1997), it is particularly difficult to address legally. Krieger (1998) adds, "Title VII is poorly equipped to control prejudice resulting from in-group favoritism. ... In-group favoritism manifests itself gradually in subtle ways. It is unlikely to trigger mobilization of civil rights remedies because instances of this form of discrimination tend to go unnoticed. If they are noticed, they will frequently seem genuinely trivial or be economically unfeasible to pursue. ... For this reason as for others, we cannot expect existing equal opportunity tools adequately to prevent, identify, or redress this more modern form of discrimination" (pp. 1325–1326). It is apparent that new techniques are needed to address this and other contemporary forms of racism.

Developing individual, intergroup, and societal-level interventions that not only control the expressions of aversive racism but also address the negative components of aversive racism has critical social implications. Aversive racism represents a latent form of bias whose expression is strongly moderated by social circumstances and norms. A change in conditions or norms can allow this bias to operate more directly and openly (see Rogers & Prentice-Dunn, 1981). For instance, research on interracial aggression has demonstrated that under normal circumstances whites are not more aggressive and harmful toward blacks than toward whites. Overt and unprovoked aggression toward blacks would readily be perceived as racist. However, when whites are first antagonized by another person's aggressiveness, when they feel freed from prevailing norms through conditions that make them feel anonymous and deindividuated, or when norms change from censuring to supporting aggression, whites exhibit more aggressiveness toward blacks than toward whites (Donnerstein & Donnerstein, 1973; Donnerstein, Donnerstein, Simon, & Ditrichs, 1972; Kawakami, Spears, & Dovidio, 2002; Rogers & Prentice-Dunn, 1981).

Latent racism also has important organizational implications. In corporate settings, racial discrimination in personnel selection decisions emerges

when norms, which typically sanction discrimination, change and an organizational authority condones discrimination (Brief, Buttram, Elliott, Reizenstein, & McCline, 1995). Thus, besides its subtle contemporary influence, if left unaddressed, aversive racism provides the seed for bias to emerge when conditions allow or encourage a more open expression of discrimination.

Future research on aversive racism has several productive avenues to pursue. One direction involves further exploration of ways to identify aversive racists. Although there is debate about the meaning and validity of implicit associations, we believe that these techniques can provide valuable insights into the dynamics of aversive racism. Karpinski and Hilton (2000) have argued that the types of responses measured by implicit techniques, such as the Implicit Association Test (Greenwald et al., 1998), reflect environmental associations rather than personal endorsement of evaluations and stereotypes. In our earlier formulation of the aversive racism framework (Gaertner & Dovidio, 1986), we hypothesized that a broad range of cognitive, affective, and motivational influences contributed to the unconscious forces involved in aversive racism. One of these was that exposure to sociocultural influences relating to institutional and cultural racism (Jones, 1997), the cultural transmission of stereotypes (Schaller, Conway, & Tanchuk, 2002), and the portrayal of blacks in the mass media (Devine, 1989) would create and support unconscious negative associations with blacks and influence behavior toward blacks in consequential ways.

The facts that (1) these associations come to mind often automatically and without intentional control for many whites, as implicit techniques have demonstrated, and (2) that these associations predict subtle behavioral manifestations of bias (Dovidio et al., 2002; Son Hing et al., 2002), particularly in the absence of conscious personal endorsement, are consistent with our aversive racism framework. In addition, although many of the techniques used to measure implicit evaluation and stereotypes focus on processes reflective of "cold" cognition, with little involvement of affective process, implicit racial evaluations are also related to fundamental affective and motivational reactions. Phelps et al. (2000) demonstrated that responses on the Implicit Association Test, but not responses on a self-report measure of prejudice, predicted activation of the amygdala, a subcortical structure of the brain that plays a role in emotional learning and evaluation and is responsive to threats. Thus, although we still contend that unconscious affective and cognitive processes contribute to aversive racism in ways beyond the effects of implicit negative associations (Nail et al., 2003), we acknowledge that implicit associations are a key component and that the techniques developed to measure implicit evaluations and stereotypes are important tools for studying aversive racism and for distinguishing aversive racists from truly nonprejudiced people.

Before the development of such measures, we generally assumed that, given the pervasive psychological and social forces promoting bias, most white Americans who said they were not prejudiced were actually aversive racists. However, evidence of the prevalence of implicit negative racial attitudes and stereotypes (Blair, 2001), the dissociation between explicit attitudes and implicit associations (Dovidio et al., 2001), and the different types of influences on behavior of implicit and explicit attitudes (Dovidio, Kawakami, & Gaertner, 2002; Fazio et al., 1995) suggest that the distinction between aversive racists and truly nonprejudiced people is a critical one that can now be examined directly.

Indeed, work on implicit attitudes has already produced evidence that people who are high on implicit prejudice exhibit particularly strong motivations to suppress negative interracial behavior (Richeson & Shelton, 2003) and are particularly sensitive to threats to their egalitarian images (Son Hing et al., 2002). Recently Son Hing, Chung-Yan, Grunfeld, Robichaud, and Zanna (2004) also demonstrated that among people who report on explicit measures that they are low in prejudice, those who have negative implicit attitudes show the subtle pattern of discrimination that characterizes aversive racism, whereas those low on implicit as well as explicit prejudice behave in a consistent egalitarian manner. In addition, understanding the developmental factors that are associated with low levels of implicit biases (Towles-Schwen & Fazio, 2001) and the types of interventions that can successfully combat implicit biases (Kawakami et al., 2000) can help to address a crucial factor in aversive racism, unconscious negative attitudes and associations.

A second potential direction for future research on aversive racism is to explore the potential evolution of social biases from overt and blatant forms to aversive forms. Norms against the expression of prejudice associated with certain types of stigmas vary considerably (Crandall & Eshleman, 2003). In general, the extent to which a person's membership in a negatively viewed out-group (i.e., a stigmatized group) is perceived to be controllable is one of the strongest determinants of whether individuals will openly express negative feelings and beliefs and discrimination (Weiner, 1995). For instance, people who possess stigmas that are perceived to be more controllable (e.g., homosexuality, obesity, alcoholism) are regarded much more negatively and are more likely to be the targets of open discrimination. However, over time some of these group memberships become more socially acceptable and sanctions against the expression of prejudice become stronger. For instance, attitudes toward homosexuals have become more favorable in recent years, and increasingly local laws and organizational policies prohibit discrimination against gays and lesbians. Thus, studying attitudes and discrimination against homosexuals may provide a case study of how prejudices evolve from blatant to aversive types. Recent research (Hebl et al., 2002) has

already demonstrated that, at least in some circumstances, although discrimination against people perceived to be gay did not occur in open, formal ways, people did discriminate in more subtle, interpersonal ways (e.g., nonverbally). Similarly, studying how subtle biases (e.g., toward people of Middle Eastern descent) become more blatant in response to current events (e.g., terrorist attacks, wars) can provide further insight into the etiology of aversive racism.

A third direction for future research involves the development of strategies and interventions to combat aversive racism more broadly, and to design strategies that will have sustained impact. As we have illustrated in this chapter, strategies that emphasize a common group identity may be effective at decreasing discrimination and changing motivations from avoiding wrongdoing, which can have inadvertent negative consequences, to motivations to do what is right. Strengthening these types of positive motivations so that they become automatically activated (Kawakami et al., 2000; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999) can counteract the effect of the influence of implicit prejudice. However, in a society in which race has a special significance historically and socially, and racism has roots in both institutions and culture (Jones, 1997), it may be difficult to sustain this common group identity and positive motivation.

Hewstone (1996) has argued that, at a practical level, interventions designed to create a common, inclusive identity may not be sufficiently potent to "overcome powerful ethnic and racial categorizations on more than a temporary basis" (p. 351). Nevertheless, creating perceptions of common group identity may form the foundation allowing other processes to operate with complementary effects (Pettigrew, 1998). For instance, creating a common ingroup identity can reduce intergroup threat and anxiety, which can increase the likelihood of intergroup contact and lead to more personalized interactions. Personalized self-disclosing interactions, in turn, further reduce intergroup bias (Brewer & Miller, 1984; Miller, 2002) and can create the kinds of experiences that are associated with low levels of implicit prejudice (Towles-Schwen & Fazio, 2001). Thus, interventions to combat aversive racism may need to involve a "cocktail" of strategies that reciprocally address the intergroup, personal, and social-cognitive roots of racial biases.

In conclusion, over the past 35 years we have argued that aversive racism is a fundamental and insidious form of racial bias that significantly affects the lives of blacks in the United States. In this chapter, we have documented the effect, examined the social and psychological processes that contribute to aversive racism, and have explored techniques to combat it. Aversive racism produces significant adverse outcomes for blacks (e.g., through bias in selection decisions, as our employment selection and college admissions studies show) and it subtly influences the nature of interracial interactions

in ways that contribute to the miscommunication and mistrust that have historically characterized race relations in the United States. Aversive racists display their biases subtly (e.g., in terms of nonverbal behaviors) and inconsistently (e.g., mainly in situations in which a negative response can be justified on the basis of a nonracial factor). As a consequence, blacks often receive conflicting messages in their interactions with aversive racists, and they perceive whites' behaviors as unpredictable and thus their positive responses as insincere and untrustworthy (Crocker & Major, 1994). Moreover, because people tend to overattribute intentionality to another person's actions (Jones & Harris, 1967), particularly for actions that are negative, blacks who feel discriminated against will likely assume that the white person's behaviors were motivated by conscious, "old-fashioned" racism. Because an aversive racist does not discriminate with conscious intention and is not aware that he or she is discriminating on the basis of race, an aversive racist will be quick to deny evidence of personal prejudice. An aversive racist's denial of intentional discrimination, although genuine, may then intensify racial conflict and distrust.

Thus, it is important that people, including both whites and blacks, become aware of the existence and impact of aversive racism to understand the different perspectives of members of different racial groups, to facilitate more effective communication and, ultimately, to take appropriate personal, social, and legal action to create a truly egalitarian society. Because aversive racists genuinely endorse egalitarian principles, once aware of their biases, they can help contribute to the solution rather than to the problem of racial tension, conflict, and inequality.

Acknowledgments

References

Abrams, D. (1985). Focus of attention in minimal intergroup discrimination. *British Journal of Social Psychology, 24*, 65–74.

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York: Harper.

Allport, G. W. (1954). *The nature of prejudice*. New York: Addison-Wesley.

Anderson, J. (1996). Black and blue. *New Yorker*, April 29 and May 6, 62–64.

Bachman, B. A. (1993). *An intergroup model of organizational mergers.* Unpublished Ph.D. Dissertation, Department of Psychology, University of Delaware, Newark, DE.

Banker, B. S. (2002). *Intergroup conflict and bias reduction in stepfamilies: A longitudinal examination of intergroup relations processes.* Unpublished Ph.D. Dissertation, Department of Psychology, University of Delaware, Newark, DE.

Banker, B. S., & Gaertner, S. L. (1998). Achieving stepfamily harmony: An intergroup relations approach. *Journal of Family Psychology, 12,* 310–325.

Bargh, J. A. (1990). Auto-motives: Preconscious determinants of thought and behavior. In E. Higgins & R. Sorrentino (Eds.), *Handbook of motivation and cognition* (Vol. 2, pp. 93–130). New York: Guilford.

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.

Blank, R. M. (2001). An overview of trends in social and economic well-being, by race. In N. J. Smelser, W. J. Wilson, & F. Mitchell (Eds.), *Racial trends and their consequences* (Vol. 1, pp. 21–39). Washington, DC: National Academy Press.

Blumer, H. (1958). Race prejudice as a sense of group position. *Pacific Sociological Review, 1,* 3–7.

Bobo, L. (1999). Prejudice as group position: Micro-foundations of a sociological approach to racism and race relations. *Journal of Social Issues, 55*(3), 445–472.

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

Bodenhausen, G. V., & Macrae, C. N. (1996). The self-regulation of intergroup perception: Mechanisms and consequences of stereotype suppression. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 227–253). New York: Guilford.

Brewer, M. B. (1979). Ingroup bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86,* 307–324.

Brewer, M. B., & Miller, N. (1984). Beyond the contact hypothesis: Theoretical perspectives on desegregation. In N. Miller & M. B. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (pp. 281–302). Orlando, FL: Academic Press.

Brief, A. P., Buttram, R. T., Elliott, J. D., Reizenstein, R. M., & McCline, R. L. (1995). Releasing the beast: A study of compliance with orders to use race as a selection criterion. *Journal of Social Issues, 51*(3), 177–193.

Brief, A. P., Dietz, J., Cohen, R. R., Pugh, S. D., & Vaslow, J. B. (2000). Just doing business: Modern racism and obedience to authority as explanations for employment discrimination. *Organizational Behavior and Human Decision Processes, 81,* 72–97.

Campbell, D. T. (1965). Ethnocentric and other altruistic motives. In D. Levine (Ed.), *Nebraska symposium on motivation* (Vol. 13, pp. 283–311). Lincoln, NE: University of Nebraska Press.

Cannon-Bowers, J. A., & Salas, E. (1999). Team performance and training in complex environments: Recent findings from applied research. *Current Directions in Psychological Science, 7,* 83–87.

Chen, M., & Bargh, J. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology, 33,* 541–560.

Clayton, S. D., & Tangri, S. S. (1989). The justice of affirmative action. In F. A. Blanchard & F. Crosby (Eds.), *Affirmative action in perspective* (pp. 177–192). New York: Springer-Verlag.

Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin, 129,* 414–446.

Crocker, J., Luhtanen, R., Broadnax, S., & Blaine, B. E. (1999). Belief in U.S. government conspiracies against blacks among black and white college students: Powerlessness or system blame? *Personality and Social Psychology Bulletin, 25,* 941–953.

Crocker, J., & Major, B. (1994). Reactions to stigma: The moderating role of justifications. In M. P. Zanna & J. M. Olson (Eds.), *The psychology of prejudice: The Ontario symposium. Ontario symposium on personality and social psychology* (Vol. 7, pp. 289–314). Hillsdale, NJ: Erlbaum.

Crocker, J., & Schwartz, I. (1985). Prejudice and ingroup favoritism in a minimal intergroup situation: Effects of self-esteem. *Personality and Social Psychology Bulletin, 11,* 379–386.

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology, 8,* 377–383.

Devine, P. G. (1989). Stereotypes and prejudice: The automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5–18.

Devine, P. G., & Monteith, M. J. (1993). The role of discrepancy-associated affect in prejudice reduction. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in intergroup perception* (pp. 317–344). Orlando, FL: Academic Press.

Donnerstein, E., & Donnerstein, M. (1973). Variables in interracial aggression: Potential ingroup censure. *Journal of Personality and Social Psychology, 27,* 143–150.

Donnerstein, E., Donnerstein, M., Simon, S., & Ditrichs, R. (1972). Variables in interracial aggression: Anonymity, expected retaliation, and a riot. *Journal of Personality and Social Psychology, 22,* 236–245.

Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology, 22,* 22–37.

Dovidio, J. F., & Gaertner, S. L. (1981). The effects of race, status and ability on helping behavior. *Social Psychology Quarterly, 44,* 192–203.

Dovidio, J. F., & Gaertner, S. L. (1983a). Race, normative structure, and help-seeking. In B. M. DePaulo, A. Nadler, & J. D. Fisher (Eds.), *New directions in helping* (Vol. 2, pp. 285–302). New York: Academic Press.

Dovidio, J. F., & Gaertner, S. L. (1983b). The effects of sex, status, and ability on helping behavior. *Journal of Applied Social Psychology, 13,* 191–205.

Dovidio, J. F., & Gaertner, S. L. (1991). Changes in the nature and assessment of racial prejudice. In H. Knopke, J. Norrell, & R. Rogers (Eds.), *Opening doors: An appraisal of race relations in contemporary America* (pp. 201–241). Tuscaloosa, AL: University of Alabama Press.

Dovidio, J. F., & Gaertner, S. L. (1996). Affirmative action, unintentional racial biases, and intergroup relations. *Journal of Social Issues, 52*(4), 51–75.

Dovidio, J. F., & Gaertner, S. L. (1998). On the nature of contemporary prejudice: The causes, consequences, and challenges of aversive racism. In J. Eberhardt & S. T. Fiske (Eds.), *Confronting racism: The problem and the response* (pp. 3–32). Newbury Park, CA: Sage.

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., Anastasio, P. A., & Sanitioso, R. (1992). Cognitive and motivational bases of bias: The implications of aversive racism for attitudes toward Hispanics. In S. Knouse, P. Rosenfeld, & A. Culbertson (Eds.), *Hispanics in the workplace* (pp. 75–106). Newbury Park, CA: Sage.

Dovidio, J. F., Gaertner, S. L, & Kawakami, K. (1998, October). *Multiple attitudes and contemporary racial bias.* Paper presented at the annual meeting of the Society for Experimental Social Psychology, Lexington, KY.

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:* Intergroup relations (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., Smith, J. K., Donnella, A. G., & Gaertner, S. L. (1997). Racial attitudes and the death penalty. *Journal of Applied Social Psychology, 27,* 1468–1487.

Duckitt, J. (1992). *The social psychology of prejudice.* Westport, CT: Praeger.

Elvira, M. M., & Zatzick, C. D. (2002). Who's displaced first? The role of race in layoff decisions *Industrial Relations, 41,* 329–361.

Esses, V. M., Dovidio, J. F., & Hodson, G. (2002). Public attitudes toward immigration in the United States and Canada in response to the September 2001 "Attack on America." *Analysis of Social Issues and Public Policy, 2,* 69–85.

Esses, V. M., Dovidio, J. F., Jackson, L. M., & Armstrong, T. M. (2001). The immigration dilemma: The role of perceived group competition, ethnic prejudice, and national identity. *Journal of Social Issue, 57,* 389–412.

Fairchild, H. H., & Cowan, G. (1997). The O. J. Simpson trial: Challenges to science and society. *Journal of Social Issues, 53*(3), 583–591.

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.

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

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology, 54,* 297–327.

Feagin, J. R., & Sikes, M. P. (1994). *Living with racism: The black middle-class experience.* Boston, MA: Beacon Press.

Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology, 73,* 31–44.

Ferguson, C. K., & Kelley, H. H. (1964). Significant factors in over-evaluation of own groups' products. *Journal of Abnormal and Social Psychology, 69,* 223–228.

Frederico, C. M., & Sidanius, J. (2002). Racism, ideology, and affirmative action revisited: The antecedents and consequences of "principled objections" to affirmative action *Journal of Personality and Social Psychology, 82,* 488–502.

Frey, D. L., & Gaertner, S. L. (1986). Helping and the avoidance of inappropriate interracial behavior: A strategy that perpetuates a nonprejudiced self-image. *Journal of Personality and Social Psychology, 50,* 1083–1090.

Gaertner, S. L. (1973). Helping behavior and racial discrimination among liberals and conservatives. *Journal of Personality and Social Psychology, 25,* 335–341.

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.

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The common ingroup identity model: Recategorization and the reduction of intergroup bias. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 4, pp. 1–26). New York: John Wiley & Sons.

Gaertner, S. L., Dovidio, J. F., Banker, B., Rust, M., Nier, J., Mottola, G., & Ward, C. (1997). Does racism necessarily mean anti-blackness? Aversive racism and pro-whiteness In M. Fine, L. Powell, L. Weis, & M. Wong (Eds.), *Off white* (pp. 167–178). London: Routledge.

Gaertner, S. L., Dovidio, J. F., Nier, J., Ward, C., & Banker, B. (1999). Across cultural divides: The value of a superordinate identity. In D. Prentice & D. Miller (Eds.), *Cultural divides: Understanding and overcoming group conflict* (pp. 173–212). New York: Russell Sage Foundation.

Gaertner, S. L., Mann, J. A., Dovidio, J. F., Murrell, A. J., & Pomare, M. (1990). How does cooperation reduce intergroup bias? *Journal of Personality and Social Psychology, 59,* 692–704.

Gaertner, S. L., Mann, J., Murrell, A., & Dovidio, J. F. (1989). Reducing intergroup bias: The benefits of recategorization. *Journal of Personality and Social Psychology, 57,* 239–249.

Gaertner, S. L., & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. *Social Psychology Quarterly, 46,* 23–30.

Gaertner, S. L., Rust, M. C., Dovidio, J. F., Bachman, B. A., & Anastasio, P. A. (1996). The Contact Hypothesis: The role of a common ingroup identity on reducing intergroup bias among majority and minority group members. In J. L. Nye & A. M. Brower (Eds.), *What's social about social cognition?* (pp. 230–360). Newbury Park, CA: Sage.

Gallup (2002). *Poll topics & trends: Race relations.* Washington, DC: The Gallup Organization. http://www.gallup.com/poll/topics/race.asp.

Greenwald, A., & Banaji, M. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27.

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.

Hamilton, D. L., & Trolier, T. K. (1986). Stereotypes and stereotyping: An overview of the cognitive approach. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 127–163). Orlando, FL: Academic Press.

Hebl, M. R., Foster, J. B., Mannix, L. M., & Dovidio, J. F. (2002). Formal and interpersonal discrimination. A field study of bias toward homosexual applicants. *Personality and Social Psychology Bulletin, 28,* 815–825.

Hewstone, M. (1996). Contact and categorization: Social psychological interventions to change intergroup relations. In C. N. Macrae, M. Hewstone, & C. Stangor (Eds.), *Foundations of stereotypes and stereotyping* (pp. 323–368). New York: Guilford.

Hodson, G., Dovidio, J. F., & Gaertner, S. L. (2002). Processes in racial discrimination: Differential weighting of conflicting information. *Personality and Social Psychology Bulletin, 28,* 460–471.

Houlette, M., Gaertner, S. L., Johnson, K. M., Banker, B. S., Riek, B. M., & Dovidio, J. F. (2004). Developing a more inclusive social identity: An elementary school intervention. *Journal of Social Issues, 60,* 35–55.

Howard, J. M., & Rothbart, M. (1980). Social categorization for in-group and out-group behavior. *Journal of Personality and Social Psychology, 38,* 301–310.

Huo, Y. J., Smith, H. H., Tyler, T. R., & Lind, A. E. (1996). Superordinate identification, subgroup identification, and justice concerns: Is separatism the problem. Is assimilation the answer? *Psychological Science, 7,* 40–45.

Hyers, L. L., & Swim, J. K. (1998). A comparison of the experiences of dominant and minority group members during an intergroup encounter. *Group Processes and Intergroup Relations, 1,* 143–163.

Insko, C. A., Schopler, J., Gaertner, L., Wildschut, T., Kozar, R., Pinter, B., Finkel, E. J., Brazil, D. M., Cecil, C. L., & Montoya, M. R. (2001). Interindividual-intergroup discontinuity reduction through the anticipation of future interaction. *Journal of Personality and Social Psychology, 80,* 95–111.

Jenkins, R. R. (2001). The health of minority children in the year 2000: The role of government programs in improving the health status of America's children. In N. J. Smelser, W. J. Wilson, & F. F. Mitchell (Eds.), *Racial trends and their consequences* (Vol. 2, pp. 351–370). Washington, DC: National Academy Press.

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.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology, 3,* 1–24.

Jones, J. M. (1997). *Prejudice and racism* (2nd ed.). New York: McGraw-Hill.

Karpinski, A., & Hilton, J. L. (2000). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81,* 774–788.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology, 55,* 893–905.

Katz, I., Wackenhut, J., & Hass, R. G. (1986). Racial ambivalence, value duality, and behavior. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 35–59). Orlando, FL: Academic Press.

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.

Kawakami, K., Spears, R., & Dovidio, J. F. (2002). Disinhibition of stereotyping: Context, prejudice and target characteristics. *European Journal of Social Psychology, 32,* 517–530.

Kelman, H. C. (2001). Reflections on social and psychological processes of legitimization and delegitimization. In J. T. Jost & B. Major (Eds.), *The psychology of legitimacy: Emerging perspectives on ideology, justice, and intergroup relations* (pp. 54–73). New York: Cambridge University Press.

Kleinpenning, G., & Hagendoorn, L. (1993). Forms of racism and the cumulative dimension. *Social Psychology Quarterly, 56,* 21–36.

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.

Kravitz, D. A. (1995). Attitudes toward affirmative action plans directed at blacks: Effects of plan and individual differences. *Journal of Applied Social Psychology, 25*, 2192–2220.

Krieger, L. H. (1995). The content of our categories: A cognitive bias approach to discrimination and equal employment opportunity. *Stanford Law Review, 47*, 1161–1248.

Krieger, L. H. (1998). Civil rights perestroika: Intergroup relations after affirmative action. *California Law Review, 86*, 1251–1333.

Lane, J. D., & Wegner, D. M. (1995). The cognitive consequences of secrecy. *Journal of Personality and Social Psychology, 69*, 237–253.

Lind, A. E., Kurtz, S., Musante, L., Walker, L., & Thibaut, J. W. (1980). Procedure and outcome effects on reactions to adjudicated resolution of conflicts of interest. *Journal of Personality and Social Psychology, 39*, 643–653.

Massey, D. S. (2001). Residential segregation and neighborhood conditions in U. S. metropolitan areas. In N. J. Smelser, W. J. Wilson, & F. Mitchell (Eds.), *Racial trends and their consequences* (Vol. 1, pp. 391–434). Washington, DC: National Academy Press.

McConahay, J. B. (1983). Modern racism and modern sexism: The effects of race, racial attitudes, and context on simulated hiring decisions. *Personality and Social Psychology Bulletin, 9*, 551–558.

McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–125). Orlando, FL: Academic Press.

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37*, 435–442.

Miller, N. (2002). Personalization and the promise of contact theory. *Journal of Social Issues, 58*, 387–410.

Monteith, M. J., Sherman, J., & Devine, P. G. (1998). Suppression as a stereotype control strategy. *Personality and Social Psychology Review, 1*, 63–82.

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.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology, 77*, 167–184.

Murphy-Berman, V. A., Berman, J. J., & Campbell, E. (1998). Factors affecting health-care allocation decisions: A case of aversive racism? *Journal of Applied Social Psychology, 28*, 2239–2253.

Murrell, A. J., Dietz-Uhler, B. L., Dovidio, J. F., Gaertner, S. L., & Drout, C. (1994). Aversive racism and resistance to affirmative action: Perceptions of justice are not necessarily color blind. *Basic and Applied Social Psychology, 15*, 71–86.

Myrdal, G. (1944). *An American dilemma: The Negro problem and modern democracy*. New York: Harper.

Nail, P. R., Harton, H. C., & Decker, B. P. (2003). Political orientation and modern versus aversive racism: Tests of Dovidio and Gaertner's (1998) Integrated Model. *Journal of Personality and Social Psychology, 84*, 754–770.

Nier, J. A. (2003). *How dissociated are implicit and explicit racial attitudes? A bogus pipeline approach*. Unpublished manuscript, Department of Psychology, Connecticut College, New London, CT.

Nier, J. A., Gaertner, S. L., Dovidio, J. F., Banker, B. S., & Ward, C. M. (2001). Changing interracial evaluations and behavior: The effects of a common group identity. *Group Processes and Intergroup Relations, 4,* 299–316.

Operario, D., & Fiske, S. T. (1998). Racism equals power plus prejudice: A social psychological equation for racial oppression. In J. Eberhardt & S. T. Fiske (Eds.), *Confronting racism: The problem and the response* (pp. 33–53). Newbury Park, CA: Sage.

Otten, S., & Moskowitz, G. B. (2000). Evidence for implicit evaluative in-group bias: Affect-based spontaneous trait inference in a minimal group paradigm. *Journal of Experimental Social Psychology, 36,* 77–89.

Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology, 42,* 1051–1068.

Pettigrew, T. F. (1959). Regional differences in anti-Negro prejudice. *Journal of Abnormal and Social Psychology, 59,* 28–36.

Pettigrew, T. F. (1998). Intergroup Contact Theory. *Annual Review of Psychology, 49,* 65–85.

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in Western Europe. *European Journal of Social Psychology, 25,* 57–76.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12,* 729–738.

Regents of the University of California v. Bakke (1978). *U. S. Law Weekly, 46,* 4896.

Richeson, J., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science, 14,* 287–290.

Robinson, P. H., & Darley, J. M. (1995). *Justice, liability, and blame.* Boulder, CO: Westview Press.

Rogers, R. W., & Prentice-Dunn, S. (1981). Deindividuation and anger-mediated interracial aggression: Unmasking regressive racism. *Journal of Personality and Social Psychology, 41,* 63–73.

Rosenfeld, R. A. (1998). Race and sex differences in career dynamics. In K. A. Myers & C. D. Anderson (Eds.), *Feminist foundations: Toward transforming sociology. Gender and society readers* (Vol. 3, pp. 113–151). Thousand Oaks, CA: Sage.

Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin, 26,* 1315–1328.

Sawires, J. N., & Peacock, M. J. (2000). Symbolic racism and voting behavior on Proposition 209. *Journal of Applied Social Psychology, 30,* 2093–2099.

Schaller, M., Conway, L. G. III, & Tanchuk, T. L. (2002). Selective pressures on the once and future contents of ethnic stereotypes: Effects of the communicability of traits. *Journal of Personality and Social Psychology, 82,* 861–877.

Schuman, H., Steeh, C., Bobo, L., & Krysan, M. (1997). *Racial attitudes in America: Trends and interpretations.* Cambridge, MA: Harvard University Press.

Sears, D. O., Henry, P. J., & Kosterman, R. (2000). Egalitarian values and contemporary racial politics. In D. O. Sears, J. Sidanius, & L. Bobo (Eds.), *Racialized politics: The debate about racism in America* (pp. 75–117). Chicago, IL: University of Chicago Press.

Shelton, J. N. (2000). A reconceptualization of how we study issues of racial prejudice. *Personality and Social Psychology Review, 4,* 374–390.

Sherif, M. (1966). *Group conflict and cooperation: Their social psychology.* London: Routledge and Kegan Paul.

Sidanius, J., Levin, S., & Pratto, F. (1998). Hierarchical group relations. Institutional terror, and the dynamics of the criminal justice system. In J. Eberhardt & S. T. Fiske (Eds.), *Confronting racism: The problem and the response* (pp. 136–165). Newbury Park, CA: Sage.

Skrentny, J. D. (1996). *The ironies of affirmative action: Politics, culture, and justice in America.* Chicago, IL: University of Chicago 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.

Smith, H. J., & Tyler, T. R. (1996). Justice and power: When will justice concerns encourage the advantaged to support policies which redistribute economic resources and the disadvantaged to willingly obey the law? *European Journal of Social Psychology, 26,* 171–200.

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., Bobocel, D. R., & Zanna, M. P. (2002). Meritocracy and opposition to affirmative action: Making concessions in the face of discrimination. *Journal of Personality and Social Psychology, 83,* 493–509.

Son Hing, L. S., Chung-Yan, G., Grunfeld, R., Robichaud, L., & Zanna, M. P. (2004). Exploring the discrepancy between implicit and explicit prejudice: A test of aversive racism theory. In J. P. Forgas, K. Williams, & W. von Hippel (Eds.). *Social Motivation: Conscious and unconscious processes* (in press).

Son Hing, L. S., Li, W., & Zanna, M. P. (2002). Inducing hypocrisy to reduce prejudicial responses among aversive racists. *Journal of Experimental Social Psychology, 38,* 71–78.

Stephan, W. G., & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues, 41*(3), 157–175.

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

Supreme Court of the United States (2003). Grutter v. Bollinger et al., Certiorrari to the United States Court of Appeals for the Sixth Circuit, *539, U.S.,* 1–5.

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues, 25*(4), 79–97.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223,* 96–102.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–48). Monterey, CA: Brooks/Cole.

Towles-Schwen, T., & Fazio, R. H. (2001). The origins of racial attitudes: Correlates of childhood experiences. *Personality and Social Psychology Bulletin, 27,* 162–175.

Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes* (Vol. 2, pp. 77–122). Greenwich, CT: JAI Press.

Turner, M. E., & Pratkanis, A. R. (1994). Affirmative action as help: A review of recipient reactions to preferential selection and affirmative action. *Basic and Applied Social Psychology, 15,* 43–69.

U.S. Merit Systems Protection Board (1997). *Fair and equitable treatment: A progress report on minority employment in the federal government.* Washington, DC: Office of Policy and Evaluation.

Vorauer, J. D., & Kumhyr, S. M. (2001). Is this about you or me? Self- versus other-directed judgments and feelings in response to intergroup interaction *Personality and Social Psychology Bulletin, 27,* 706–719.

Whaley, A. L. (1998). Racism in the provision of mental health services: A social-cognitive analysis. *American Journal of Orthopsychiatry, 68,* 47–57.

Wegner, D. (1994). Ironic processes of mental control. *Psychological Review, 101,* 34–52.

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct.* New York: Guilford.

Wilder, D. A. (1981). Perceiving persons as a group: Categorization and intergroup relations. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 213–257). Hillsdale, NJ: Erlbaum.

Williams, R. M., Jr. (1947). *The reduction of intergroup tensions.* New York: Social Science Research Council.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107,* 101–126.

Wittenbrink, B., Judd, C., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72,* 262–274.

Worchel, S., Rothgerber, H., Day, E. A., Hart, D., & Butemeyer, J. (1998). Social identity and individual productivity with groups. *British Journal of Social Psychology, 37,* 389–413.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology, 10,* 109–120.

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology, 73,* 73–90.

Wyer, N., & Hamilton, D. (1998). The balance between excitation and inhibition in stereotype use. In Wyer (Ed.), *Stereotype activation and inhibition. Advances in social cognition* (Vol. 11, pp. 227–242). Mahwah, NJ: Erlbaum.