12 Aversive Racism and Contemporary Bias

John F. Dovidio, Samuel L. Gaertner, and Adam R. Pearson

In the United States, the 1960s and early 1970s were characterized by significant societal changes. The Civil Rights Movement and social, political, and moral forces stimulated these changes to address racism by White Americans toward Black Americans and achieve the nation's historical egalitarian ideals. With the Civil Rights legislation and other federal mandates, it was no longer simply immoral to discriminate against Blacks; it was now also illegal. Surveys and national polls revealed significant reductions in overt expressions of prejudice among Whites toward Blacks (Dovidio & Gartner, 2004). This unprecedented change in race relations in the United States changed the nature of racial attitudes, from blatant to subtle, and consequently the study of prejudice in psychology (Dovidio, 2001). In other countries, similar normative changes have reduced blatant expressions of prejudice while more subtle, yet equally pernicious, forms of bias persist (see Pettigrew & Meertens, 1995).

This chapter reviews the development of theory about contemporary forms of racism – focusing primarily on aversive racism – tracing the evolution of this perspective, describing key empirical evidence, and identifying productive avenues for future research. We begin by reviewing relations among different theories of subtle contemporary racism, including aversive racism, and then discuss work on implicit prejudice and its relationship to aversive racism. We then consider the implications of aversive racism for interventions to reduce bias and identify promising new directions for research on contemporary racism, in general, and aversive racism, in particular.

Overview of Theories of Subtle Racism

The changing social norms and values shaped by the civil rights era posed unique challenges to the study of prejudice. Although overt expressions of prejudice and negative stereotyping have substantially declined, in part because of new normative pressures toward egalitarianism, privately held beliefs continue to reflect negative racial attitudes and beliefs. One effect of these new norms was that people appeared to more deliberately manage how others perceived their racial attitudes. For example, when expressing attitudes

under conditions in which they were led to believe their true attitudes could be detected (e.g., bogus pipeline; Roese & Jamieson, 1993), Whites displayed significantly more negative attitudes toward Blacks than when they reported their attitudes under more normal conditions. This was, in part, because of the conscious management of the expression of prejudice and self-reports of prejudice and interracial behaviors, particularly those that were performed privately or were not likely to be perceived as blatantly racist diverged.

Changes in law and societal norms toward the affirmation of egalitarian values in the 1960s also prompted another important change – a shift in the nature of racial prejudice in the United States. Many Whites not only altered their public attitudes to appear less prejudiced but also developed a private self-concept of being non-prejudiced. They genuinely embraced these egalitarian norms and internalized them. With prejudice increasingly recognized as undesirable, for many, consciously holding racist attitudes or beliefs was incompatible with a positive self-image. These changes to the expression and self-control of racial attitudes, along with methodological advances in capturing subtle and unintended manifestations of bias, spawned a new era of research on race relations, beginning in the 1970s. We briefly describe four historically influential approaches to understanding subtle racism dating from this period: symbolic, modern, ambivalent, and aversive forms of racism.

Symbolic Racism

Symbolic racism theory (Sears & Henry, 2005) developed in response to a practical problem: the failure of traditional self-report measures to predict people's responses to racially targeted policies and Black political candidates. Sears, Henry, and Kosterman (2000) observed, "Few Whites now support the core notions of old-fashioned racism . . . Our own view is that the acceptance of formal equality is genuine but that racial animus has not gone away; it has just changed its principal manifestations" (p. 77). Symbolic racism involves four basic belief components that reflect the confluence of politically conservative, individualistic values and early acquired negative racial affect that typically emerge in adolescence before many other sociopolitical beliefs. These include (a) discrimination against Blacks is "a thing of the past,"(b) Blacks' failure to progress is attributable to their unwillingness to work hard enough, (c) Blacks make excessive demands, and (d) Blacks have gotten more than they deserve. The components are typically measured using self-reported responses (see Henry & Sears, 2002).

Symbolic racism predicts people's political attitudes and behavior better than do measures of old-fashioned racism, realistic threats, and perceived intergroup competition, non-racial ideologies (e.g., individualism, egalitarianism), and political affiliation and ideology (Sears & Henry, 2005). Specifically, symbolic racism uniquely predicts White Americans' attitudes toward a range of racially relevant policies, including busing for school integration and affirmative action, as well as

less explicitly race-targeted policies that disproportionately affect blacks, such as policies relating to crime and welfare. Symbolic racism also predicts opposition to Black political candidates, as well as support for ethnocentric White candidates, such as former Ku Klux Klan leader David Duke (Tesler & Sears, 2010).

Modern Racism

Modern racism theory was derived from symbolic racism theory; however, the two positions diverge on the hypothesized origins of bias. Whereas symbolic racism proposes that Whites' negative attitudes are primarily rooted in concerns that Blacks threaten Whites' worldviews by violating principles of individualism, modern racism theory hypothesizes that various forms of negative affect (e.g., fear, disgust), which may be acquired through early socialization and modeling, persist into adulthood. Both theories, though, assume that these feelings are expressed indirectly and symbolically, in terms of more abstract social and political issues (e.g., opposition to school integration, ostensibly to support neighborhood schools; see McConahay, 1986).

Modern racism is assessed using the Modern Racism Scale, a self-report measure similar to that used to assess symbolic racism. The scale was originally designed to be an indirect measure of racism that is less susceptible to social desirability concerns (McConahay, 1986). Like symbolic racism, modern racism predicts voting against political candidates who are Black or sympathetic toward Blacks, and voting against policies designed to assist Blacks such as affirmative action and school integration programs. It also predicts these political attitudes better than do measures of political conservatism, identification as a Democrat or Republican, education, and most importantly, personal interests in the outcomes of a vote (Henry, 2009).

The modern and symbolic racism approaches have been challenged on similar conceptual grounds. Because the scales are not pure, direct measures of stereotypes or prejudice against Blacks, critics have argued that both modern racism and symbolic racism scales tap non-racial principles underlying conservatism (such as opposition to excessive government intervention) rather than racial attitudes, per se, and there is evidence that a subset of high modern racism scorers are, indeed, principled conservatives rather than racists (Son Hing, Chung-Yan, Hamilton, & Zanna, 2008). Other critiques of the Modern Racism Scale have contended that this measure should no longer be classified as an indirect measure but, rather, as a blatant measure of racism. Indeed, responses on the Modern Racism Scale appear to be more susceptible to social desirability influences and self-presentational concerns than in the past (Calanchini & Sherman, 2013). As McConahay (1986) noted 30 years ago, "It is expected that new items will have to be generated for the Modern Racism Scale as new issues emerge in American race relations and some of the current items become more reactive while the ambivalence lingers" (p. 123). Nevertheless, items from the Modern Racism Scale still predict responses to political issues (e.g., evaluations of Obama versus McCain during the 2008

presidential campaign, support for government policies benefiting Blacks) in ways independent of measures of blatant prejudice toward Blacks (Ditonto, Lau, & Sears, 2013).

Ambivalent Racism

Another important psychological insight into contemporary forms of prejudice and discrimination involved the recognition of the ambivalence that many Whites hold toward Blacks. Pioneering work by Katz (1981) proposed the existence of Whites' ambivalent attitudes toward Blacks. Katz (1981; see also Katz, Wackenhut, & Hass, 1986) noted that attitudes toward Blacks often involve sympathy and a desire to be fair and egalitarian, as well as negative affective reactions (Dovidio & Gaertner, 2004; Gaertner & Dovidio, 1986; Katz, 1981). These inconsistent and opposing elements create psychological tension that amplifies responses to the associated groups and their members (Katz, 1981). Response amplification involves responding more intensely, positively or negatively, to members of stigmatized groups than non-stigmatized groups as a consequence of the psychological tension involved in some forms of bias (e.g., toward Blacks and people with disabilities). For instance, when participants were induced to be overly harsh in their evaluation of a Black person (compared to giving mild negative feedback or a harsh evaluation of a White person), they compensated more strongly by subsequently helping the Black person more (including more than the harshly evaluated White person (Katz, Glass, Lucido, & Farber, 1979). However, when White participants responded overly negatively to a confederate (administering noxious vs. non-offensive noise) and were not given the opportunity to compensate for their harsh action by helping the confederate in a subsequent task, they denigrated the Black confederate uniquely strongly (Katz, Glass, & Cohen, 1973). These and other findings by Katz and his colleagues (see Katz et al., 1986) provide general support for the concept of ambivalence-amplification.

Aversive Racism

Whereas modern and symbolic racism characterize the attitudes of political conservatives, aversive racism characterizes the biases of those who are politically liberal (Nail, Harton, & Becker, 2003) and believe that they are not prejudiced, but whose unconscious negative feelings and beliefs get expressed in subtle, indirect, and often rationalizable ways. In 1970, Kovel 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 [sic] represents the open flame of racial hatred" (p. 54). Building on Kovel's (1970) distinction, over the past 45 years we have explored the existence and operation of aversive racism among White Americans. Aversive racism is hypothesized to be qualitatively different from blatant, old-fashioned, racism. Aversive racists sympathize with victims of past

injustice, support principles of racial equality, and genuinely regard themselves as non-prejudiced, but at the same time they possess conflicting, often noncon-scious, negative feelings and beliefs about Blacks that are rooted in basic psychological processes (e.g., social categorization) that promote racial bias (Dovidio & Gaertner, 2004; Gaertner & Dovidio, 1986; Pearson, Dovidio, & Gaertner, 2009). In addition, the negative feelings that aversive racists have toward Blacks do not reflect open hostility or hatred. Instead, aversive racists' reactions typically involve discomfort, anxiety, or fear. That is, while they find Blacks aversive, they find any suggestion that they might be prejudiced aversive as well. In addition, in some instances aversive racism can reflect the expression of more positive feelings toward Blacks than toward Whites (Gaertner et al., 1997; Greenwald & Pettigrew, 2014). The experience of differential positivity rather than negativity toward racial ingroup and outgroup members similarly can obscure self-recognition of prejudicial attitudes.

Like symbolic and modern racism, aversive racism is hypothesized to operate in subtle and indirect ways. However, unlike symbolic and modern racism approaches that utilize self-report scales to measure these concepts and have largely focused on their relationship to political behavior, work on aversive racism has examined a broad array of responses to Blacks and Whites. Because the basic premises of aversive racism are not tied to specific scale items and the U.S. political context, the principles of aversive racism are applicable to behaviors of dominant groups toward minorities in other nations that also have strong societal egalitarian values, such as Canada (Son Hing et al., 2008), England (Hodson, Hooper, Dovidio, & Gaertner, 2005), the Netherlands (Kleinpenning & Hagendoorn, 1993), Portugal (de França & Monteiro, 2013), and Spain (Wojcieszak, 2015).

The aversive racism framework also helps identify when discrimination against Blacks and other minority groups will occur. Because aversive racists consciously endorse egalitarian values and truly believe that they are non-prejudiced, they will not act inappropriately when discrimination would be obvious to others and themselves. Specifically, when they are presented with a situation in which the normatively appropriate response is clear, 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. However, discrimination will occur in situations in which normative structure is weak, when the guidelines for appropriate behavior are vague, when the basis for social judgment is ambiguous, or when one 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 aversive racists to maintain their self-image as non-prejudiced.

Support for the aversive racism framework has been obtained across a broad range of paradigms and participant populations. Early tests of the aversive racism framework focused on prosocial behavior for both theoretical and practical

reasons. Theoretically, because aversive racists are hypothesized to be particularly effective at censoring negative behavior toward Blacks, the biases associated with aversive racism may often manifest as differential prosocial responses toward Whites and Blacks in need. Indeed, it was research on the differential behavior of Whites toward Black and White motorists who were stranded on a highway that represented the first empirical work on aversive racism (Gaertner, 1973). The results of a meta-analysis of 31 experiments on Whites' interracial helping behavior conducted over the past 40 years evidence a stable pattern of discrimination reflective of aversive racism that has not subsided over time (Saucier, Miller, & Doucet, 2005). Based on these findings, the authors concluded that racism and discrimination against Blacks "can and will exist as long as individuals harbor negativity toward Blacks at the implicit level" (p. 14).

Aversive racism is also influential in selection decisions in employment and college admission, interpersonal judgments, and policy and legal decisions (see Dovidio & Gaertner, 2004; Knight, Guiliano, & Sanchez-Ross, 2001; Sommers & Ellsworth, 2000). For example, Dovidio and Gaertner (2000) examined White college students' support for hiring Black and White applicants for a selective campus position within the same college in the years 1989 and 1999. When the candidates' credentials clearly qualified or disqualified them for the position, there was no discrimination against the Black candidate (i.e., the highly qualified Black candidate was just as likely to be hired as the highly qualified White candidate). However, when candidates' qualifications for the position were less obvious and the appropriate decision was more ambiguous (moderate qualifications), White participants recommended a Black candidate significantly less often than a White candidate with exactly the same credentials. Whereas overt expressions of prejudice (measured by items on a self-report scale) declined over this 10-year period, the pattern of subtle discrimination in selection decisions remained essentially unchanged.

Additional research offers further insight into processes that underlie these effects. When ambiguous or mixed credentials are involved, people systematically weigh credentials differently based on their unconscious biases. For example, when providing input to college admission decisions for candidates with mixed credentials (e.g., strong high school grades but modest standardized scores, or vice versa), White college students emphasized the credential that White candidates were stronger in relative to Black candidates as being the more valid predictor of success in college. This differential weighting of the credentials, in turn, justified students' stronger recommendations of White than Black candidates for admission (Hodson, Dovidio, & Gaertner, 2002; see also Brief, Dietz, Cohen, Pugh, and Vaslow, 2000, for employment bias against Blacks; Rooth, 2007, for hiring biases against Muslims).

Recent developmental research is also supportive of the hypothesized role of egalitarian norms in the dynamics of aversive racism. Developmentally, older children are more aware of egalitarian norms than are younger children, and thus they may be more likely to show patterns consistent with aversive racism,

whereas younger children may display more blatant intergroup bias. For instance, de França and Monteiro (2013, Study 1) found that younger White Portuguese children (ages 5–7) generally favored other White children over Black children in the distribution of rewards (candy) regardless of how they performed on a task. Older children (ages 8–10) rewarded relatively poorly performing Black children less than White children but, consistent with the aversive racism perspective, did not discriminated against Black relative to White children when they performed equally well. In another study (de França & Monteiro, 2013, Study 1), increasing the salience of egalitarian norms through the presence of an adult reduced bias in helping a White relative to a Black child for the older children but not for younger children.

Further support for this developmental time-course was obtained by McGillicuddy-De Lisi, Daly, and Neal (2006), who found evidence of aversive racism in the way that 10-year-olds distributed resources to Black and White story characters. Whereas younger children (7- to 8-year-olds) distributed money equivalently, regardless of race, age, performance, or need, older children (9- to 10-year-olds) showed evidence of differential allocation by race, consistent with aversive racism. Older children allocated less money to needy Black than to needy White children but allocated more money to Black than to White children who performed unusually successfully on the designated task. Furthermore, whereas the younger children justified their allocations based on principles of equality for both Black and White targets, the older children made different justifications for their allocations for White and Black targets, appealing to principles of equality for the former and equity (i.e., deservingness) for the latter. Together, these findings implicate a window of middle-to-late childhood for the emergence of cognitive and behavioral strategies that allow aversive racists to establish and maintain a non-prejudiced self-concept while engaging in subtle discriminatory practices.

More generally, the studies cited show that, in contrast to the dramatic decline in overt expressions of prejudice, subtle forms of discrimination persist and, despite their subtlety, can be as consequential and pernicious as more overt forms. In addition to the outcomes detailed earlier in the context of hiring and helping behavior, even minor discriminatory patterns can have substantial effects over time. For instance, computer simulations of corporate hiring decisions demonstrate that even a slight systematic bias (a 1% bias against women in performance assessments) compounds to produce a substantial underrepresentation of women in top-level management positions (Martell, Lane, & Emrich, 1996), which approximates the actual gender disparity in top-level management positions today (35% women vs. 65% men).

Early work on aversive racism had limited means to assess people's hypothesized unconscious attitudes and beliefs. Major developments in assessing implicit biases over the past several decades, however, have offered a solution to this problem (see Chapter 11 by Yogeeswaran, Devos, and Nash

for a full discussion of implicit bias and measures). Whereas self-report methods today are often used to assess explicit attitudes, implicit attitudes can be assessed through a wide range of methodologies, including response latency tasks, memory tasks, and physiological measures (e.g., heart rate and galvanic skin response), as well as indirect self-report measures (e.g., biases in behavioral and trait attributions). In the next section, we discuss research on implicit attitudes and its relevance to aversive racism.

Aversive Racism and Implicit Bias

Among the most common techniques for assessing implicit, automatically activated (and potentially unconscious) attitudes are response latency techniques (Fazio & Olson, 2003; Gaertner & McLaughlin, 1983). As Yogeeswaran et al. review (Chapter 11, this title), response time techniques have been commonly used to test the strength of association between social groups and positive or negative valence (implicit attitudes) and particular traits or beliefs (implicit stereotypes), as well as the extent to which subgroups (e.g., Blacks) are perceived as belonging within a superordinate group (e.g., Americans).

One of the most popular of the response time techniques in this area is the Implicit Association Test (IAT; see Greenwald, Poehlman, Uhlmann, & Banaji, 2009). The IAT relies on the basic assumption that people are faster categorizing groups of stimuli that are similar in valence than groups of stimuli that are dissimilar in valence. For instance, research using the Race-IAT has shown that people who associate negativity with Blacks are quicker to respond when images of Blacks and unpleasant words share a response key than when images of Blacks and pleasant words share a response key (Greenwald et al., 2009). Importantly, responses on these types of measures are difficult to control and are often unintended. Implicit attitudes are acquired early in life (Rudman, Phelan, & Heppen, 2007) from repeated exposure to positive or negative information about a group, or a lack of exposure, either through socialization or direct experience, and can be resistant to change in response to new information (Gregg, Seibt, & Banaji, 2006; Smith & DeCoster, 2000; see also Yogeeswaran et al.'s chapter in this title).

Consistent with the aversive racism framework, in a large sample of more than 14,000 U.S. Whites, whereas most (59.4%) appeared non-prejudiced on a self-report (explicit) measure, a large majority (71.5%) nevertheless showed evidence of implicit bias on the Race-IAT that was largely dissociated from their explicit (self-reported) views (Greenwald & Krieger, 2006). Moreover, the percentage of respondents who display pro-White implicit race bias varies relatively little by age, sex, political ideology, and educational attainment (Greenwald & Krieger, 2006). For instance, more liberal individuals display less explicit prejudice on self-report measures but show pro-White implicit racial bias at levels equal to those of conservatives. Also, whereas younger children exhibit racial bias both explicitly

and implicitly, as they grow older and more responsive to egalitarian norms, explicit bias declines but implicit bias remains high (Baron & Banaji, 2006; Dunham, Baron, & Banaji, 2008).

Initially, research on implicit attitudes focused on developing and refining measurement techniques, distinguishing implicit from explicit measures, and clarifying the origins and meaning of implicit measures of attitudes. More recent work has examined the predictive validity of implicit measures. A meta-analysis by Greenwald et al. (2009; cf. Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013) of 122 research reports found that both implicit measures of attitudes such as the IAT and explicit measures predicted a range of behaviors to attitude objects (explicit attitudes average $r = .36$ and implicit attitudes average $r = .27$). However, the predictive validity of explicit measures for socially sensitive issues was much weaker. Across a range of racial issues in particular (e.g., willingness to hire a Black job applicant [Ziegert & Hanges, 2005], subjective and neural threat responses to Blacks [Phelps et al., 2000], affective and nonverbal expressions toward Blacks [McConnell & Leibold, 2001]), implicit attitudes (average $r = .24$) were a better predictor overall than explicit attitudes (average $r = .12$).

Notably, different theoretical perspectives suggest that a key factor in the relative validity of implicit and explicit measures for predicting behavior is the context in which the behavior occurs and the type of behavior being examined (Dovidio et al., 2001; Dovidio, Kawakami, Smoak, & Dovidio, 2009). For example, Fazio's (1990) MODE model indicates that whereas implicit measures will better predict spontaneous behaviors, explicit measures will better predict deliberative behaviors, including those in situations in which social desirability factors are salient (Fazio & Olson, 2003). Accordingly, the relative influence of explicit and implicit attitudes depends on the type of response that is made. Whereas explicit attitudes shape more deliberative responses in which the costs and benefits of various courses of action are weighed, implicit attitudes influence uncontrollable behaviors and responses that people do not view as indicative of their attitude.

Consistent with this distinction, and paralleling the findings in racial contexts in the United States, in a Canadian sample, Son Hing et al. (2008) found that when assessing candidates with more moderate qualifications, evaluators recommended White candidates more strongly for a position than Asian candidates with identical credentials. However, when evaluating candidates with exceptionally strong qualifications, no such selection bias emerged. Moreover, the researchers found that implicit bias against Asians (as measured by an IAT), but not explicit prejudice, predicted weaker support for hiring Asian candidates who had moderate qualifications. However, when the Asian candidate had distinctively strong qualifications, neither implicit nor explicit prejudice predicted the hiring decision.

Divergent effects of implicit and explicit racial attitudes can also fuel divergent perspectives and experiences of Blacks and Whites in interracial interactions. Dovidio, Kawakami, and Gaertner (2002) demonstrated that whereas Whites' explicit (self-reported) racial attitudes predicted their relatively controllable verbal expressions in their interactions with Blacks, Whites' implicit attitudes, which

were generally negative, predicted their nonverbal behaviors. Moreover, Black interaction partners weighed the nonverbal behavior more heavily than the verbal behavior in their impressions of the White partner and the interaction. Thus, Whites and Blacks had divergent perspectives in their interactions, and Blacks' awareness of conflicting positive verbal and negative nonverbal behavior undermined how trustworthy they saw the White interaction partner.

Stigmatization and discrimination can have profound effects on the health of members of disadvantaged groups (see Paradies, Bastos, and Priest, Chapter 25, this title). In addition, more specifically, implicit biases of physicians can have a substantial impact on the quality of health care that Blacks receive and ultimately affect their health and well-being. For example, White physicians generally see themselves as non-prejudiced and color blind but also harbor negative implicit racial biases toward Blacks (Sabin, Rivara, & Greenwald, 2008). Physicians' implicit biases predict medical recommendations, such as lower quality of coronary care for Black patients (Green et al., 2007), in ways independent of explicit racial bias. The effects of implicit bias may be particularly influential for medical issues that may evoke stereotypical beliefs about Blacks (e.g., illegal drug use). For instance, physicians higher in implicit bias are substantially less willing to prescribe narcotics to ease the pain of Black patients than White patients (Sabin & Greenwald, 2012). Also, consistent with research on the influence of implicit racial bias in social interactions (Dovidio et al., 2002), doctors higher in implicit bias speak faster to and have shorter visits with Black patients (Cooper et al., 2012), and they display less warmth in their medical interactions (Penner et al., 2010). Overall, physicians higher in implicit bias are less patient centered in their care of Black patients (Blair et al., 2013; Cooper et al., 2012).

Moreover, as suggested by the aversive racism framework, the combination of explicit and implicit attitudes is also an important determinant of physician-patient relations. Penner et al. (2010) found that although Black patients generally perceived doctors higher in implicit racial bias less favorably, this effect was particularly pronounced when doctors were low in explicit prejudice but high in implicit bias – the aversive racist profile. Specifically, Black patients who interacted with aversive racist physicians were less satisfied with the interaction and felt less close to their physicians than Black patients who interacted with other physicians, including physicians who were high on both implicit and explicit bias. These effects may be due, in part, to a lack of awareness of personal bias among implicitly biased physicians. As in Dovidio et al.'s (2002) study, the perceptions of doctors themselves were based primarily on their explicit racial attitudes: Doctors higher in explicit bias reported that they involved Black patients less in the medical decision-making process during the visit.

Taken together, current research reveals that implicit biases contribute in significant ways to racism. They contribute to disparate treatment of minorities in ways independent of explicit (self-reported) personal biases.

Implications for Interventions

Prejudice-reduction techniques have traditionally been concerned with changing conscious attitudes (explicit attitudes) and obvious expressions of bias, and they have commonly utilized educational programs aimed at combating such views and behaviors (Stephan & Stephan, 2001). However, antibias education programs often prove to be ineffective in producing intended outcomes in organizations (e.g., Kalev, Dobbin, & Kelly, 2006) and educational settings (Yeager & Walton, 2011), as well as among the general public (see Paluck & Green, 2009). Contemporary forms of racism, such as aversive racism, can be more difficult to combat than more blatant forms because implicit bias is more difficult to detect than overt bias (even for low-status group members), and high-status group members are often less attuned to cues of subtle discrimination than are members of low-status groups (Salvatore & Shelton, 2007). Aversive racists already recognize prejudice as harmful, but they do not recognize that *they* are prejudiced. Other techniques are thus required. In this section, we consider different types of interventions for reducing subtle bias.

Reducing Implicit Bias

To the extent that most Whites in the United States endorse egalitarian values, one obvious strategy to combat contemporary bias is to tailor interventions to minimize or eliminate implicit bias. Indeed, new techniques have been developed to modify implicit biases and their expression (Lai et al., 2014). These techniques include approaches that encourage people to think of members of other groups more in terms of their individual qualities than as primarily members of their racial or ethnic groups. Examples of these techniques include training in perspective taking and empathic responding or increasing exposure to counter-stereotypic members of the group. However, implicit attitudes are "habits of mind" (Devine, 1989) that may be deeply entrenched, and thus difficult, if not impossible, to unlearn completely. As a consequence, although the various techniques identified by Lai et al. (2014) may temporarily inhibit the activation of implicit biases in a given context, the effects may not be durable. To the extent that media, social models, and other societal influences reinforce stereotypes and justify group hierarchy (Jost et al., 2012; Sidanius & Pratto, 1999), implicit biases can become reactivated beyond the lab.

In addition, interventions to address implicit bias through suppression can have paradoxical effects. Implicit biases are activated automatically without utilizing cognitive resources. Because conscious efforts to suppress implicit biases are cognitively taxing, implicit biases may be especially pronounced and influential (i.e., producing a "rebound effect"; Monteith, Sherman, & Devine, 1998) when those cognitive resources are depleted and the ability to control one's behavior consciously is diminished (Trawalter & Richeson, 2006). Implicit cognition plays

a particularly important role in shaping behavior when cognitive control is impaired, for example, by fatigue or tasks that strain cognitive resources. Attempts to suppress implicit bias can also affect impressions of others. Because suppression can make implicit biases hyperaccessible due to rebound, these thoughts can color the attributions of Black partners in social interactions, leading to more negative impressions of them, often in ways that justify negative reactions to them (Pearson, Dovidio, Phills, & Onyeador, 2014). Also, emphasizing the automatic nature of implicit bias can lead people to believe that their biases are fixed rather than malleable, which may reduce motivations to change (Neel & Shapiro, 2012).

Although bias control has traditionally been conceptualized in terms of conscious efforts to inhibit negative attitudes and stereotypes that become activated in one's mind, studies suggest that biases may also be combated at the implicit level through non-conscious processes that inhibit their activation in the first place (Calanchini, Gonsalkorale, Sherman, & Klauer, 2013). For instance, individuals with chronic explicit race-related egalitarian goals spon-taneously activate egalitarian thoughts when primed with Black (vs. White) faces, whereas individuals who do not report such goals do not (Moskowitz, Salomon, & Taylor, 2000). In addition, individuals with chronic egalitarian goals do not exhibit the automatic activation of racial stereotypes that Whites typically demonstrate when exposed to Blacks (Park, Glaser, & Knowles, 2008). Activating egalitarian goals even temporarily, for example, by having people describe a personal incident in which they failed to be egalitarian toward Blacks, can subsequently inhibit non-conscious racial stereotyping (see Moskowitz & Ignarri, 2009, for a review).

In general, whereas conscious efforts to avoid stereotyping may often fail or even exacerbate bias because individuals lack insight into the processes that promote and regulate it, implicit egalitarian goals may succeed by co-opting the very psychological mechanisms that sustain it, replacing stereotypic associations with egalitarian or atypical associations when perceiving or interacting with mem-bers of other racial and ethnic groups. However, developing implicit egalitarian goals may often require extended practice to establish (Devine, Forscher, Austin, & Cox, 2012).

Correcting for Unconscious Bias

Rather than attempting to alter implicit biases, another approach involves inter-vening to inhibit the unfair consequences of this bias through conscious control. Many Whites believe that unfair bias against Blacks is no longer prevalent and that anti-White bias is a bigger societal problem than anti-Black bias (Norton & Sommers, 2011). In addition, aversive racists who do acknowledge that prejudice exists and is harmful do not recognize that they are prejudiced.

A foundational step in combating the influence of subtle bias in admissions decisions is thus educating those making decisions about the existence of implicit

prejudice and subtle bias and their personal susceptibility to these influences. The decision process therefore needs to be structured in ways that facilitate the detection of bias. For instance, because people can rationalize preferences in complex decisions in many ways, they can readily justify decisions about a given individual. However, bias that is subtle in a particular case becomes much more apparent when its cumulative effects are considered. Thus, one particularly effective way for addressing subtle bias is to assess the outcomes of a set of decisions in terms of whether the desired diversity among candidates has been achieved (Crosby, Clayton, Alksnis, & Hemker, 1986). In addition, reminding individuals of their personal accountability for detecting and correcting for unconscious bias promotes fairer, less biased decision making in employment contexts (Nadler, Lowery, Grebinoski, & Jones, 2014). Because implicit biases affect decision making subtly and unintentionally, good intentions are not sufficient to eliminate this bias. Instead, the emphasis of diversity efforts needs to recognize diversity early in the decision-making process and in assessing the outcomes.

One implication of this approach is that it is important that people acknowledge race and its potential effects on their behavior, rather than attempting to be color blind and dismiss the influence of race on their perceptions, cognitions, feelings, and actions. Whites are typically motivated to avoid seeing themselves as racially biased and often adopt a color-blind strategy when engaging in interracial interactions, particularly when they anticipate racial tension (Apfelbaum, Norton, & Sommers, 2012). However, efforts to be color blind can sometimes produce rebound effects, causing biases to become activated even more strongly. Indeed, Uhlmann and Cohen (2005) found that participants who were more confident in the objectivity of their judgments were also more likely to discriminate against equally qualified female candidates for a stereotypically male job (chief of police), inflating criteria that favored male over female candidates. Ironically, the act of affirming a non-prejudiced self-image can further increase the likelihood that even ostensibly non-prejudiced individuals will discriminate (see also Effron, Cameron, & Monin, 2009; Monin & Miller, 2001).

Whites' attempts to be color blind can also alienate minority group members, who generally seek acknowledgment of their racial identity and further contribute to interracial distrust. Apfelbaum, Sommers, and Norton (2008) found that although avoidance of race was seen as a favorable strategy by Whites for promoting more positive interracial interactions, in practice failure to acknowledge race actually predicted decrements in Whites' nonverbal friendliness and resulted in greater perceptions of racial prejudice by Black interaction partners. By contrast, a multicultural perspective, which involves the appreciation of differences and common connections, represents a promotion focus (Scheepers, Saguy, Dovidio, & Gaertner, 2014) that encourages mutual understanding and cooperation. When members of different groups adopt a multicultural orientation in their interaction, they generally show greater

responsiveness to one another's goals, attitudes, and needs, and the interaction is more productive and mutually satisfying (Vorauer, Gagnon, & Sasaki, 2009).

Harnessing Egalitarian Motives

It may also be possible to capitalize on aversive racists' good intentions and induce self-motivated efforts to reduce the impact of unconscious biases by making them aware of these biases. Son Hing, Li, and Zanna (2002) examined responses of people identified as non-prejudiced (low in both explicit and implicit prejudice) and aversive racists (low in explicit prejudice but high in implicit prejudice) to self-awareness of one's own hypocrisy. In a study conducted in Canada with Asians as the target minority group, participants were assigned to either a hypocrisy condition, in which they reflected on situations in which they had reacted negatively or unfairly toward an Asian person, or to a control condition, in which they were not asked to write about such situations. The researchers predicted that making people aware of violations of their egalitarian principles would arouse guilt among aversive racists (who harbor negative feelings toward Asians) and thus produce compensatory behavior when recommending funding for Asian student groups among aversive racists but not among non-prejudiced participants. The results supported the predictions. 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, however, 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.

Work by Monteith and colleagues (see Monteith, Arthur, & Flynn, 2010, for a review) further indicates that when low-prejudiced people recognize discrepancies between their behavior (i.e., what they *would* do) and their personal standards (i.e., what they *should* do) toward minorities, they feel guilt and compunction, which subsequently produces motivations to respond without prejudice in the future. With practice, these individuals learn to reduce prejudicial responses and respond in ways that are consistent with their non-prejudiced personal standards. When extended over time, this process of self-regulation can produce sustained changes in even automatic negative responses.

Dovidio, Kawakami, and Gaertner (2000) found that greater discrepancies between what one would do and should do produced higher levels of guilt among Whites in an initial experimental session, and this relationship occurred primarily for low-prejudiced participants. These findings indicate the potential recruitment of self-regulatory processes for low-prejudiced but not high-prejudiced participants. When participants returned three weeks later, there was generally greater alignment (i.e., smaller discrepancy) between what one would and should do – an indication that both high-prejudiced and low-prejudiced participants showed a decrease in

overt expressions of bias. However, both groups differed in terms of the extent to which they internalized these changes. Low-prejudiced Whites who had larger initial discrepancies showed greater reductions in implicit stereotyping; in contrast, for high-prejudiced Whites the relationship was weaker and nonsignificant.

Nevertheless, feedback suggesting successful self-regulation in some cases may backfire. Mann and Kawakami (2012) explored the effect of goal feedback on outgroup discrimination and ingroup favoritism, as well as implicit racial attitudes. In a series of studies, participants showed greater racial bias after receiving feedback that they were progressing on egalitarian goals compared to feedback that they were failing on egalitarian goals or a no-feedback condition. Specifically, participants who were told that they were progressively becoming more egalitarian subsequently sat farther away from Blacks and closer to Whites and demonstrated greater implicit racial prejudice. Thus, feedback that signals reduced need to monitor and regulate one's responses may undermine Whites' egalitarian motives. Nevertheless, Whites who are both aware of their biases and remain concerned about displaying biases show a high level of motivation to engage in activities (e.g., training) to eliminate their racial bias over time (Perry, Murphy, & Dovidio, 2015).

Taken together, these findings demonstrate that the good intentions of aversive racists can be harnessed to promote self-initiated change in both conscious and unconscious biases with sufficient awareness, effort, and practice.

Redirecting the Forces of Ingroup Bias

According to the aversive racism framework, the negative attitudes and feelings that develop toward members of other groups are rooted in basic socio-cognitive processes. One such process is the categorization of people into ingroups and outgroups (see Dovidio & Gaertner, 2010, for a review). Because categorization is a basic process fundamental to intergroup bias, this process has been targeted in efforts to combat the negative effects of aversive racism.

The common ingroup identity model (Gaertner & Dovidio, 2000, 2012) is one such intervention approach that harnesses social categorization as a means to reduce intergroup bias and has received strong empirical support in interventions with both child and adult populations. Specifically, if members of different groups are induced to think of themselves as a single superordinate ingroup rather than as two separate groups, attitudes toward former outgroup members will become more positive by reaping the benefits of ingroup status. Enhancing the salience of a common ingroup identity has been shown to inhibit the activation of both implicit (Van Bavel & Cunningham, 2009) and explicit (Gaertner & Dovidio, 2000) biases. Thus, by changing the basis of categorization from race to an alternative, inclusive dimension, one can alter who "we" are and who "they" are, undermining a potent contributing force to contemporary racism. The formation of a common identity, however, need not require groups to forsake their subgroup identities. It is possible for members to

conceive of themselves as holding a "dual identity" in which other identities and a superordinate identity are salient simultaneously (Crisp, Stone, & Hall, 2006; Gaertner & Dovidio, 2009).

This intervention was used in a primary care setting with Black patients being seen by non-Black physicians (Penner et al., 2013). Compared to a standard-of-care control condition, Black patients who were induced to feel a common identity and team goal (successful resolution of their medical problem) had more successful interactions. Black patients in the common identity condition, compared to a control condition, reported more trust of their physician four weeks later. This greater trust led to greater subsequent adherence to the physician's treatment recommendations.

Future Directions

The past decade has seen substantial progress in understanding manifestations and expressions of subtle bias, and particularly in the development of techniques for assessing implicit forms of bias, incorporating perspectives from affective and cognitive neuroscience and psychophysiological research. Next, we consider three domains that have received comparatively less attention and represent particularly promising areas for future research: (a) conditions under which implicit biases may be expressed blatantly, (b) the relationship between contemporary bias and intimate relations, and (c) extensions of the dynamics of aversive racism to other groups.

From Subtle to Overt Bias: The Inversion Problem

Research described in this chapter suggests that "everyday" prejudice, bias within the latitude of normal expression, provides a foundation that, under appropriate conditions, can be manifested in actions that systematically disadvantage members of other groups. This unfair treatment can sometimes escalate to physical harm. Indeed, history is replete with examples, from Rwanda, the former Yugoslavia, and the Holocaust during World War II, demonstrating how neighbors who had been living in apparent harmony can suddenly become violent enemies. Modern-day examples also abound, such as the recent rise in blatant homophobia in Russia and Africa and Islamophobia in the United States and Europe. Moreover, events, such as the terrorist attack of 9/11 can create strong negative associations with a group (Muslims) that was not strongly stigmatized previously (Esses, Dovidio, & Hodson, 2002).

How and when do aversive biases become overt? Because aggression and intergroup violence are usually normatively sanctioned, aversive racists may be particularly inhibited in engaging in interracial aggression in most contexts. Nevertheless, given their underlying negative feelings and beliefs, aversive racists

may also be especially susceptible to disinhibiting influences that can lead to overt hostility and aggression. Moreover, as Stephan and Stephan (2001) propose, inter-racial anxiety can amplify Whites' affective reactions and consequently produce more extreme behavioral responses to Blacks than to Whites. Thus, the more diffuse emotions of anxiety and discomfort that are experienced by aversive racists and typically lead to avoidance can readily be transformed into more intense negative emotions that motivate aggression toward Blacks (see Dovidio, Pearson, Gaertner, & Hodson, 2008). Mullen (1986), for instance, found that interracial violence by Whites against Blacks often occurs within a social context that permits or encourages aggression. His analysis of newspaper reports of Blacks being lynched by White mobs revealed that violence against Blacks was more likely when Whites were part of a larger group and experienced greater anonymity and deindividuation (see also Leader, Mullen, & Abrams, 2007). Thus, factors that normally disinhibit aggressive behavior, such as provocation and anonymity, may be especially potent within intergroup contexts for promoting interracial aggres-sion toward Blacks.

Laws and legal processes may also facilitate progression to more overt forms of discrimination. Research on aversive racism has shown that disparate treatment is most likely to occur in combination with other factors that provide non-racial justifications for negative treatment, and that disparate treatment often represents ingroup favoritism (pro-White responses) rather than outgroup antipathy (anti-Black responses; Gaertner et al., 1997; Greenwald & Pettigrew, 2014). However, when the normative context supports it, these dynamics can lead to direct physical harm. In the context of the government-sanctioned death penalty, Blacks who are perceived more stereotypically are more likely to receive the death penalty in capital cases (Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006); in addition, greater implicit dehumanization of Blacks predicts greater violence by police officers toward black children suspected of crimes (Goff, Jackson, Di Leone, Culotta, & DoTomasso, 2014). Unconscious biases can influence how evidence is perceived and weighed, affecting assessments of guilt, and once a defendant is judged guilty, they can affect the severity of recommended punishment in ways that reinforce stereotypes of Blacks as criminals and fuel dehumanization and distrust.

Contemporary Bias and Intimate Relationships

Close relationships represent an important new context for studying aversive racism, given the uniquely intimate nature of these exchanges and the well-documented power of cross-group friendships and romantic relationships to combat bias (Davies, Wright, Aron, & Comeau, 2013; Page-Gould, Mendoza-Denton, & Tropp, 2008). One in three Americans (35%) indicates that a member of his or her immediate family or a close relative is currently married to someone of a different race (Wang, 2012). Processes that help form and maintain social bonds

in interpersonal relationships may, therefore, also be essential for developing cross-group friendships and intimate relationships that can promote positive intergroup attitudes as well as egalitarian norms in social networks through extended contact processes (see Davies et al., 2013).

"Big data" approaches, analyzing survey responses and behaviors of more than 30 million online daters – nearly one-third of the U.S. population under age 40 – offer a unique window into actual dating habits of a large and growing segment of the American public (see Rudder, 2014a; and Curington, Ken-Hou, & Lundquist, 2015). Cross-sectional analyses of online daters' self-reported preferences show increasing openness to interracial dating and interracial marriage. However, online daters' behaviors (e.g., private attractiveness ratings, who communicates with and dates whom) show evidence of substantial and in some cases intensified racial bias, with White men and women rating Blacks as less attractive, and dating Blacks less often, than those of other racial and ethnic groups (Rudder, 2014b; and Curington et al., 2015, for similar findings showing strong preferences for White versus non-White among multiethnic online datees).

These effects are mirrored in national marriage statistics. In 2010, 17% of U.S. Blacks and 26% of U.S. Hispanics married outside of their race, whereas only 9% of Whites entered interracial marriages (Wang, 2012). Among Blacks and Hispanics, newlyweds who married Whites tended to have higher educational attainment than did those who married within their own racial or ethnic group.

Aversive Racism beyond Black-White Relations

Aversive and other contemporary forms of racism are hypothesized to result from the conflict between persistent negative feelings or beliefs and social norms inhibiting bias against Blacks. However, norms concerning egalitarian treatment vary substantially across groups (Crandall, Eshleman, & O'Brien, 2002), as do legal restrictions. In the United States, as of 2015, federal laws prohibit employment discrimination on the basis of age, race, ethnicity, disability, and religion, but not on the basis of weight, economic status, gender identity, or sexual orientation. At the state level, 52% of the U.S. lesbian, gay, bisexual, and transgender (LGBT) population currently live in states that do not prohibit employment discrimination based on sexual orientation or gender identity (see www.lgbtmap.org). Thus, whereas subtle bias may characterize responses to groups with legal protections, blatant biases may be commonly expressed toward groups, such as obese or poor people, who are not protected and perceived to be responsible for their disadvantage. Nevertheless, changing public opinion and legal decisions, such as the 2015 U.S. Supreme Court decision to support federal and state recognition of same-sex marriage, can signal new egalitarians norms for LGBT individuals. As a consequence, bias that was previously largely blatant toward gay men and lesbians may now more commonly be expressed in more subtle and rationalizable forms – an aversive form of bias.

Beyond the extension to different groups and types of intergroup relations, future research on aversive racism might productively consider responses to groups with multiple or intersectional identities. Traditionally, race in the United States was socially defined in strict binary terms. "One drop of Black blood" defined a person as Black rather than White, and multiracial individuals were treated as Blacks (*hypodescent*; Ho, Sidanius, Levin, & Banaji, 2011). In part stimulated by the prominence of President Barack Obama, the category multiracial has become more widely recognized and applied, one that may in some ways be more appealing than either White or Black. Nevertheless, notions of hypodescent persist in contemporary society in ways that can impact everyday relationships and broader societal intergroup relations. Curington et al. (2015), for instance, examined messages sent between heterosexual men and women between 2003 and 2010 on a large U.S. dating site and found that multiracial men and women are increasingly preferred above all other groups, including Whites. Nevertheless, persistent pro-White and anti-Black biases in dating preferences were found, with online daters generally preferring White multiethnic indivi-duals to non-Whites.

In addition to categorization based on race, members of racial groups may also be categorized simultaneously on other dimensions, such as gender. Considering differentiated responses to people with intersectional identities can help provide a more nuanced and comprehensive understanding of contemporary bias. For example, because Black men are viewed as more prototypical of the category Black than are Black women, they may be the primary target under conditions that elicit subtle bias, while Black women's "intersectional invisibility" may insulate them from this bias (Purdie-Vaughns & Eibach, 2008). However, when egalitarian norms are salient, Black women may benefit less from compensatory actions or policies than Black men. Thus, when multiple or intersectional identities are involved, subtle bias may be even more complex than research on this topic has previously considered.

Additional research might identify processes that generalize across different stigmatized groups that may contribute to the "spread" of contemporary forms of bias within and between social networks. For instance, people are more tolerant of discrimination against minorities after overhearing racist jokes (Ford & Ferguson, 2004) and exposure to cues as subtle as an antiracism t-shirt can influence uncon-scious bias (Lun, Sinclair, Glenn, & Whitchurch, 2007).

Research on prejudice has typically not examined social influence beyond dyadic interactions in the laboratory; however, emerging research on social net-works may offer some clues. In a large field experiment in U.S. high schools, Paluck (2011) trained student leaders (Peer Trainers) to confront expressions of prejudice in five randomly assigned high schools across a period of five months and compared their responses to those of students awaiting training in five control schools. Whereas the treatment Peer Trainers' egalitarian attitudes (assessed through self-report) spread only to close friends five months post-training, their anti-prejudice behavior (signing a gay rights petition) spread to both friends and

acquaintances, including to individuals outside of the Peer Trainers' school net-work. These findings suggest that conformity of non-prejudiced behavior may precede the internalization of egalitarian attitudes among more peripherally con-nected individuals. A comprehensive understanding of both structural (e.g., rela-tionship networks beyond dyads) as well as psychological factors can, thus, help guide the development of interventions that may more effectively combat con-temporary forms of bias and its transmission.

Conclusion

In virtually all societies, across time and cultures, groups are organized hierarchically, with some groups consistently enjoying greater privilege and resources than others. Racism operates to maintain the power and status of domi-nant groups over other groups. However, as a function of social, political, and historical forces, racism is manifested in different ways. In the later 1930s through the 1950s, stimulated politically by the Nazis' rise to power in Germany, histori-cally by the Holocaust, and intellectually by the classic work on the authoritarian personality (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950), racial biases were seen as not simply disruptions in rational processes but also as dangerous aberrations from normal thinking. Stimulated further by the civil rights movement in the 1960s, traditional blatant expressions of prejudice rapidly declined in the United States, but subtle and indirect forms – symbolic, modern, ambivalent, and aversive – emerged.

Aversive racism, the main focus of this chapter, represents a distinct form of contemporary bias that characterizes the racial attitudes of well-intentioned, typically political liberal people, who genuinely endorse egalitarian values and believe that they are not prejudiced. The recent and future directions in aversive racism research, which we have identified in this chapter, help broaden the theoretical perspective both practically and conceptually. Practically, the dynamics of aversive racism are not restricted to White Americans' responses to Black Americans but to other traditionally disadvantaged groups, and this phenomenon occurs cross-nationally. It also affects formal and consequential interactions and outcome related to legal processes, employment, and health care. Theoretically, the study of aversive racism contributes to understanding structural factors (e.g., color-blind policies), intergroup relations (e.g., group mistrust and conflict), interpersonal interaction (e.g., nonverbal communication), and intrapersonal emotional (e.g., emotion regulation) and cognitive (implicit cognition) processes.

The particular challenge of studying contemporary forms of bias, such as aversive racism, is not simply in identifying the basic psychological processes that underlie bias, which may be common across societies and groups, but also the unique social and political factors that may systematically alter the ways

that contemporary biases manifest. These forces may increase the subtlety of the expression of bias or, in some cases, alter the normative context in ways that justify open discrimination. Furthermore, understanding and communicating information about contemporary bias can alter the nature of racism and shape the emergence of new forms of bias. As Gergen (1973) observed more than four decades ago, "The dissemination of psychological knowledge modifies the patterns of behavior on which the knowledge is based" (p. 309). Thus, making people aware of how their behavior can reflect subtle prejudice or discrimination may, at least for some individuals, motivate them to express negative racial attitudes in even more indirect and less obvious ways, to insulate them from the recognition by others or themselves of their racial bias. Although Gergen's observation suggests the potential futility in studying contemporary racism – the study of subtle bias may produce new, "ultra-subtle" forms of prejudice and discrimination – understanding the dynamics of contemporary bias is ultimately necessary to address bias at its psychological roots and thereby contribute to more equitable and sustainable societies that benefit all groups.

References

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

Apfelbaum, E. P., Norton, M. I., & Sommers, S. R. (2012). Racial color blindness: Emergence, practice, and implications. *Current Directions in Psychological Science*, *21*, 205–209.

Apfelbaum, E. P., Sommers, S. R., & Norton, M. I. (2008). Seeing race and seeming racist? Evaluating strategic colorblindness in social interaction. *Journal of Personality and Social Psychology*, *95*, 918–932.

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6, 10 & adulthood. *Psychological Science*, *17*, 53–58.

Blair, I. V., Steiner, J. F., Fairclough, D. L., Hanratty, R., Price, D. W., Hirsh, H. K. . . . Havranek, E. P. (2013). Clinicians' implicit ethnic/racial bias and perceptions of care among Black and Latino patients. *Annals of Family Medicine*, *11*, 43–52.

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.

Calanchini, J., Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2013). Counter-prejudicial training reduces activation of biased associations and enhances response monitoring. *European Journal of Social Psychology*, *43*, 321–325.

Calanchini, J., & Sherman, J. W. (2013). Implicit processes reflect associative, non-associative, and non-attitudinal processes. *Social and Personality Psychology Compass*, *7*, 654–667.

Cooper, L. A., Roter, D. L., Carson, K. A., Beach, M. C., Sabin, J. A., Greenwald, A. G., & Inui, T. S. (2012). The associations of clinicians' implicit attitudes about race with medical visit communication and patient ratings of interpersonal care. *American Journal of Public Health*, *102*, 979–987.

Crandall, C. S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology*, *82*, 359–378.

Crisp, R. J., Stone, C. H., & Hall, N. R. (2006). Recategorization and subgroup identification: Predicting and preventing threats from common ingroups. *Personality and Social Psychology Bulletin*, *32*, 230–243.

Crosby, F., Clayton, S., Alksnis, O., & Hemker, K. (1986). Cognitive biases in the perception of discrimination: The importance of format. *Sex Roles*, *14*, 637–646.

Curington, C. V., Ken-Hou, L., & Lundquist, J. H. (2015). Positioning multiraciality in cyberspace: Treatment of multiracial daters in an online dating website. *American Sociological Review*, *80*, 764–788.

de França, D. X., & Monteiro, M. B. (2013). Social norms and the expression of prejudice: The development of aversive racism in children. *European Journal of Social Psychology*, *43*, 263–271.

Davies, K., Wright, S. C., Aron, A., & Comeau, J. (2013). Intergroup contact through friendship: Intimacy and norms. In G. Hodson & M. Hewstone (Eds.), *Advances in intergroup contact* (pp. 200–230). New York: Psychology Press.

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

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, *48*, 1267–1278.

Ditonto, T. M., Lau, R. R., & Sears, D. (2013). AMPing racial attitudes: Comparing the power of explicit and implicit racism measures in 2008. *Political Psychology*, *34*, 487–510.

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

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

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

Dovidio, J. F., & Gaertner, S. L. (2010). Intergroup bias. In S. T. Fiske, D. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (Vol. *2*, 5th ed., pp. 1084–1121). New York: Wiley

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. F., Kawakami, K., Smoak, N., & Gaertner, S. L. (2009). The nature of contemporary racial prejudice: Insights from implicit and explicit measures of

attitudes. In R. E. Petty, R. H. Fazio, & P. Brinol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 165–192). New York: Psychology Press.

Dovidio, J. F., Pearson, A. R., Gaertner, S. L., & Hodson, G. (2008). On the nature of contemporary prejudice: From subtle bias to severe consequences. In V. M. Esses & R. A. Vernon (Eds.), *Explaining the breakdown of ethnic relations: Why neighbors kill* (pp. 41-60). Malden, MA: Blackwell.

Dunham, Y., Baron, A. S., & Banaji, M. R. (2008). The development of implicit intergroup cognition. *Trends in Cognitive Sciences*, *12*, 248–253.

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science*, *17*, 383–386.

Effron, D. A., Cameron, J. S., & Monin, B. (2009). Endorsing Obama licenses favoring Whites. *Journal of Experimental Social Psychology*, *45*, 590–593.

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.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE Model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. *23*, pp. 75–109). Orlando, FL: Academic Press.

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

Ford, T. E., & Ferguson, M. A. (2004). Social consequences of disparagement humor: A prejudice norm theory. *Personality and Social Psychology Review*, *8*, 79−94.

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. (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. (2009). A Common Ingroup Identity: A categorization-based approach for reducing intergroup bias. In T. Nelson (Ed.), *Handbook of prejudice* (pp. 489–506). Philadelphia, PA: Taylor & Francis.

Gaertner, S. L., & Dovidio, J. F. (2012). Reducing intergroup bias: The Common Ingroup Identity Model. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (Vol. *2*, pp. 439–457). Thousand Oaks, CA: Sage.

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

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

Gergen, K. J. (1973). Social psychology as history. *Journal of Personality and Social Psychology*, *26*, 309–320.

Goff, P. A., Jackson, M. C., Di Leone, B. A. L., Culotta, C. M., & DoTomasso, N. A. (2014). The essence of innocence: Consequences of dehumanizing Black children. *Journal of Personality and Social Psychology*, *106*, 526–545.

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

Greenwald, A. G., & Krieger, L. H. (2006). Implicit bias: Scientific foundations. *California Law Review*, *94*, 945–967.

Greenwald, A. G., & Pettigrew, T. F. (2014). With malice toward none: Ingroup favoritism enables discrimination. *American Psychologist*, *69*, 669–684.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*, 17–41.

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, *90*, 1–20.

Henry, P. J. (2009). Modern racism. In J. M. Levine & M. A. Hogg (Eds.), *Encyclopedia of group processes and intergroup relations* (pp. 575–577). Thousand Oaks, CA: Sage.

Henry, P. J., & Sears, D. O. (2002). The Symbolic Racism 2000 Scale. *Political Psychology*, *23*, 253–283.

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.

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

Ho, A., Sidanius, J., Levin, D. T., & Banaji, M. R. (2011). Evidence for hypodescent and racial hierarchy in the categorization and perception of biracial individuals. *Journal of Personality and Social Psychology*, *100*, 492–506.

Jost, J. T., Liviatan, I., van der Toorn, J., Ledgerwood, A., Mandisodza, A., & Nosek, B. A. (2012). System justification: A motivational process with implications for social conflict. In E. Kals & J. Maes (Eds.), *Justice and conflicts: Theoretical and empirical contributions* (pp. 315–327). New York: Springer.

Kalev, A., Dobbin, F., & Kelly, E. (2006). Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review*, *71*, 589–617.

Katz, I. (1981). *Stigma: A social psychological analysis*. Hillsdale, NJ: Erlbaum.

Katz, I., Glass, D. C., & Cohen, S. (1973). Ambivalence, guilt, and scapegoating of minority group victims. *Journal of Experimental Social Psychology*, *9*, 423–436.

Katz, I., Glass, D. C., Lucido, D., & Farber, J. (1979). Harm-doing and victim's racial or orthopedic stigma as determinants of helping behavior. *Journal of Personality*, *47*, 340–364.

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.

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.

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E., Joy-Gaba, J. A. . . . Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, *143*, 1765–1785.

Leader, T., Mullen, B., & Abrams, D. (2007). Without mercy: The immediate impact of group size on lynch mob atrocity. *Personality and Social Psychology Bulletin*, *33*, 1340–1352.

Lun, J., Sinclair, S., Glenn, C., & Whitchurch, E. (2007). (Why) do I think what you think? Epistemic social tuning and implicit prejudice. *Journal of Personality and Social Psychology*, *93*, 957−972.

Mann, N. H., & Kawakami, K. (2012). The long, steep path to equality: Progressing on egalitarian goals. *Journal of Experimental Psychology: General*, *141*, 187–197.

Martell, R. F., Lane, D. M., & Emrich, C. E. (1996). Male-female differences: A computer simulation. *American Psychologist*, *51*, 157–158.

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. 99–125). 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.

McGillicuddy-De Lisi, A. V., Daly, M., & Neal, A. (2006). Children's distributive justice judgments: Aversive racism in Euro-American children? *Child Development*, *77*, 1063–1080.

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, *81*, 33–43.

Monteith, M. J., Arthur, S. A., & Flynn, S. M. (2010). Self-regulation and bias. In J. F. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.). *SAGE handbook of prejudice, stereotyping, and discrimination* (pp. 493–507). London: Sage.

Monteith, M. J., Sherman, J. W., & Devine, P. G. (1998). Suppression as a stereotype control strategy. *Personality and Social Psychology Review*, *2*, 63–82.

Moskowitz, G. B., & Ignarri, C. (2009). Implicit volition and stereotype control. *European Review of Social Psychology*, *20*, 97–145.

Moskowitz, G. B., Salomon, A. R., & Taylor, C. M. (2000). Preconsciously controlling stereotyping: Implicitly activated egalitarian goals prevent the activation of stereotypes. *Social Cognition*, *18*, 151–177.

Mullen, B. (1986). Atrocity as a function of lynch mob composition: A self-attention perspective. *Personality and Social Psychology Bulletin*, *12*, 187–197.

Nadler, J. T., Lowery, M. R., Grebinski, J., & Jones, R. G. (2014). Aversive discrimination in employment interviews: Reducing effects of sexual orientation bias with accountability. *Psychology of Sexual Orientation and Gender Diversity*, *1*, 480–488.

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.

Neel, R., & Shapiro, J. R. (2012). Is race bias malleable? Whites' lay theories of racial bias predict divergent strategies for interracial interactions. *Journal of Personality and Social Psychology*, *103*, 101–120.

Norton, M. I., & Sommers, S. R. (2011). Whites see racism as a zero-sum game that they are now losing. *Perspectives on Psychological Science*, *6*, 215–218

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, *105*, 171–192.

Page-Gould, E., Mendoza-Denton, R., & Tropp, L. R. (2008). With a little help from my cross-group friend: Reducing anxiety in intergroup contexts through cross-group friendship. *Journal of Personality and Social Psychology*, *95*, 1080–1094.

Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A critical look at evidence from the field and the laboratory. *Annual Review of Psychology*, *60*, 339–367.

Paluck, E. L. (2011). Peer pressure against prejudice: A high school field experiment examining social network change. *Journal of Experimental Social Psychology*, *47*, 350–358.

Park, S. H., Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice moderates the effect of cognitive depletion on unintended discrimination. *Social Cognition*, *26*, 401–419.

Pearson, A. R., Dovidio, J. F., & Gaertner, S. L. (2009). The nature of contemporary prejudice: Insights from aversive racism. *Social and Personality Psychology Compass*, *3*, 314–338.

Pearson, A. R., Dovidio, J. F., Phills, C. E., & Onyeador, I. N. (2014). Attitude-goal correspondence and interracial interaction: Implications for executive function and impression formation. *Journal of Experimental Social Psychology*, *49*, 907–914.

Penner, L. A., Dovidio, J. F., West, T. W., Gaertner, S. L., Albrecht, T. L., Dailey, R. K., & Markova, T. (2010). Aversive racism and medical interactions with Black patients: A field study. *Journal of Experimental Social Psychology*, *46*, 436–440.

Penner, L. A., Gaertner, S. L., Dovidio, J. F., Hagiwara, N., Porcerelli, J., Markova, T., & Albrecht, T. L. (2013). A social psychological approach to improving the outcomes of racially discordant medical interactions. *Journal of General Internal Medicine*, *28*, 1143–1149.

Perry, S. P., Murphy, M. C., & Dovidio, J. F. (2015). Modern prejudice: Subtle, but Unconscious? The role of bias awareness in Whites' perceptions of personal and others' biases. *Journal of Experimental Social Psychology*, *61*, 64–78.

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in Western Europe. *European Journal of Social Psychology*, *25*, 57–75.

Phelps, E. A., O'Conner, 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*, 1–10.

Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional invisibility: The distinctive advantages and disadvantages of multiple subordinate-group identities. *Sex Roles*, *59*, 377–391.

Roese, N. J., & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. *Psychological Bulletin*, *114*, 363–375.

Rooth, D. (2007). Automatic associations and discrimination in hiring: Real world evidence. *Labour Economics*, *17*, 523–534.

Rudder, C. (2014a). *Dataclysm: Who we are (when we think no one's looking)*. New York: Crown Publishers.

Rudder, C. (2014b). *Race and attraction, 2009–2014*. Retrieved from http://oktrends .okcupid.com/

Rudman, L. A., Phelan, J. E., & Heppen, J. B. (2007). Developmental sources of implicit attitudes. *Personality and Social Psychology Bulletin*, *33*, 1700–1713.

Sabin, J. A., & Greenwald, A. G. (2012). The influence of implicit bias on treatment recommendations for 4 common pediatric conditions: Pain, urinary tract infection, attention deficit hyperactivity disorder, and asthma. *American Journal of Public Health*, *102*, 988–995.

Sabin, J. A., Rivara, F. P., & Greenwald, A. G. (2008). Physician implicit attitudes and stereotypes about race and quality of medical care. *Medical Care*, *46*, 678–685.

Salvatore, J., & Shelton, J. N. (2007). Cognitive costs of exposure to racial prejudice. *Psychological Science*, *18*, 810–815.

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

Scheepers, D., Saguy, T., Dovidio, J. F., & Gaertner, S. L. (2014). A shared dual identity promotes cardiovascular challenge response during interethnic interaction. *Group Processes & Intergroup Relations*, *17*, 324–341.

Sears, D. O., & Henry, P. J. (2005). Over thirty years later: A contemporary look at symbolic racism. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. *37*, pp. 95–150). San Diego, CA: Academic Press.

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

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

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108–131.

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.

Son Hing, L. S., Chung-Yan, G., Hamilton, L., & Zanna, M. (2008). A two-dimensional model that employs explicit and implicit attitudes to characterize prejudice. *Journal of Personality and Social Psychology*, *94*, 971–987.

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. (2001). *Improving intergroup relations*. Thousand Oaks, CA: Sage.

Tesler, M., & Sears, D. O. (2010). *Obama's race: The 2008 election and the dream of a post-racial America*. Chicago: University of Chicago Press.

Trawalter, S., & Richeson, J. A. (2006). Regulatory focus and executive function after interracial interactions. *Journal of Experimental Social Psychology*, *42*, 406–412.

Uhlmann, E. L., & Cohen, G. L. (2005). Constructed criteria: Redefining merit to justify discrimination. *Psychological Science*, *16*, 474–480.

Van Bavel, J. J., & Cunnigham, W. A. (2009). Self-categorization with a novel mixed-race group moderates automatic social and racial biases. *Personality and Social Psychology Bulletin*, *35*, 321–335.

Vorauer, J. D., Gagnon, A., & Sasaki, S. J. (2009). Salient intergroup ideology and inter-group interaction. *Psychological Science*, *20*, 838–845.

Wang, W. (2012). *The rise of intermarriage: Rates, characteristics vary by race and gender*. Retrieved from www.pewsocialtrends.org/2012/02/16/the-rise-of-intermarriage/2/#chapter-1-overview

Wojcieszak, M. (2015). Aversive racism in Spain – Testing the theory. *International Journal of Public Opinion Research*, *27*, 22–45.

Yeager, D. S. & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*, 267–301.

Ziegert, J. C., & Hanges, P. J. (2005). Employment discrimination: The role of implicit attitudes, motivation, and a climate for racial bias. *Journal of Applied Psychology*, *90*, 554–562.