Journal of Personality and Social Psychology
2006, Vol. 91, No. 4, 652–661

0022-3514/06/$12.00 DOI: 10.1037/0022-3514.91.4.652

Stereotyping and Evaluation in Implicit Race Bias: Evidence for Independent Constructs and Unique Effects on Behavior

David M. Amodio
New York University

Patricia G. Devine
University of Wisconsin—Madison

Implicit stereotyping and prejudice often appear as a single process in behavior, yet functional neuroanatomy suggests that they arise from fundamentally distinct substrates associated with semantic versus affective memory systems. On the basis of this research, the authors propose that implicit stereotyping reflects cognitive processes and should predict instrumental behaviors such as judgments and impression formation, whereas implicit evaluation reflects affective processes and should predict consummatory behaviors, such as interpersonal preferences and social distance. Study 1 showed the independence of participants' levels of implicit stereotyping and evaluation. Studies 2 and 3 showed the unique effects of implicit stereotyping and evaluation on self-reported and behavioral responses to African Americans using double-dissociation designs. Implications for construct validity, theory development, and research design are discussed.

Keywords: prejudice, stereotyping, implicit evaluation, affect, cognition

The distinction between affect and cognition in the human psyche dates back to the earliest philosophers of the mind and continues to be a major feature of modern psychology and neuroscience. Indeed, contemporary theorists have argued that the affective–cognitive distinction is essential for understanding the mind, brain, and behavior (Cacioppo, Gardner, & Berntson, 1999; Damasio, 1994; Zajonc, 1980), and neuroscientists have delineated distinct neural pathways for basic affective and cognitive systems of learning and memory (Davis & Whalen, 2001; Squire & Zola, 1996). In the intergroup relations literature, affect and cognition traditionally correspond to two key components of race bias: prejudice and stereotyping (Allport, 1954; Devine, 1989; Dovidio, Brigham, Johnson, & Gaertner, 1996; Fiske, 1998; Mackie & Smith, 1998). Whereas the term *prejudice* refers to negative affective responses toward outgroup members (McConahay & Hough, 1976), the term *stereotype* refers to cognitive representations of culturally held beliefs about outgroup members (Hamilton, 1981).

In research on traditional, *explicit* race biases, the conceptual distinction between prejudice and stereotyping has provided a useful framework for examining their respective contributions to different forms of discrimination (Dovidio, Esses, Beach, & Gaertner, 2004; Park & Judd, 2005). By contrast, in research on more automatic, or *implicit*, forms of race bias, little attention has been given to the affective–cognitive distinction or the important implications it may have for understanding the relationship between implicit biases and behavior. Because the distinction between affective and semantic neural circuits are most pronounced in more basic, implicit levels of processing, theories of implicit race bias have much to gain by considering the alternative roles of affect and cognition.

In the present research, we examined the roles of affect and cognition in implicit race bias and their effects on behavior. On the basis of theory and research from social psychology and neuroscience, we argue for a conceptual distinction between implicit stereotyping and implicit evaluative race bias and propose that these two forms of implicit race bias are predictive of different types of discriminatory responses.

Relationship Between Implicit Stereotyping and Evaluation

A survey of the implicit race bias literature reveals that very few studies have directly examined the relation between affective and cognitive aspects of implicit bias (for reviews, see Blair, 2001; Fazio & Olson, 2003), and none have sought to obtain independent measures of implicit stereotyping versus evaluation.[1] Indeed, most expressions of race bias reflect a combination of affective and cognitive processes, and the most commonly reported African American stereotypes are negative in valence (e.g., unintelligent, hostile, poor, lazy, and dishonest; Devine & Elliot, 1995). Yet despite the common concurrence of negative valence and stereotypes of stigmatized groups, underlying distinctions between affective and cognitive components may be important for understanding mechanisms of implicit race biases and their effects on behavior.

David M. Amodio, Department of Psychology, New York University; Patricia G. Devine, Department of Psychology, University of Wisconsin—Madison.

We gratefully acknowledge Carolyn Stahlhut, Ryan Beld, and Marissa Langhoff for their assistance in data collection.

Correspondence concerning this article should be addressed to David M. Amodio, Department of Psychology, New York University, 6 Washington Place, New York, NY 10003. E-mail: David.amodio@nyu.edu

[1] Throughout this article, we use the term *implicit evaluation* rather than *implicit prejudice* as a more precise label to refer to automatic evaluative associations. By using implicit evaluation, we avoid invoking unintended connotations associated with the complicated construct of prejudice, such as consciously endorsed racist attitudes and beliefs (Amodio et al., 2003; Devine et al., 2002).

Although the distinction is seldom made, past research has featured measures that may be characterized as assessing either implicit stereotyping (e.g., Lepore & Brown, 1997; Spencer, Fein, Wolfe, Fong, & Dunn, 1998), implicit evaluation (e.g., Amodio, Harmon-Jones, & Devine, 2003; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998), or some combination of stereotyping and evaluation (e.g., Dovidio, Evans, & Tyler, 1986; Kawakami, Dion, & Dovidio, 1998; Rudman, Ashmore, & Gary, 2001; Wittenbrink, Judd, & Park, 1997, 2001). The use of such measures suggests that both are valid constructs that have been studied somewhat independently and that both forms of implicit bias are prevalent among European Americans, such that African Americans are typically associated with negative evaluations and with the culturally defined stereotype content (Blair, 2001). However, although the distinction between implicit stereotyping and implicit evaluation has been acknowledged in past work (e.g., Greenwald & Banaji, 1995; Greenwald et al., 2002), theorizing has not been advanced to directly address the relation between cognitive and affective mechanisms underlying these two forms of implicit race bias.

Distinct Neural Substrates for Basic Affective and Semantic Associations

In the neuroscience literature, neural substrates of affective forms of learning and memory have been distinguished from semantic forms, and this distinction has implications for the present set of issues. Results from decades of research on animals and humans suggest that the amygdala and its associated subcortical circuits are central to affective learning and memory (Lang, Bradley, & Cuthbert, 1990; LeDoux, 2000). This body of work has shown that affective associations are learned quickly, often after a single presentation of an unconditioned stimulus in a fear-learning paradigm. Once learned, such associations extinguish slowly, and subsequent reconditioning to the stimulus is facilitated (Bouton, 1994). It is important to note that amygdala-based learning does not depend on semantic associations; for example, mice easily learn affective associations despite their inability to process semantic information. By comparison, semantic learning and memory (e.g., conceptual priming) are embedded in mechanisms for language, believed to be supported by a phylogenetically newer network of neocortical structures that are significantly expanded among humans compared with those of other species (Gabrieli, 1998; Rissman, Eliassen, & Blumstein, 2003; Squire & Zola, 1996). Semantic associations may be learned in the absence of affective content, such that patients with a damaged amygdala retain normal semantic associations despite the loss of conditioned physiological responses in a fear-conditioning paradigm (Bechara, Damasio, & Damasio, 2003).

An examination of anatomical and neurochemical connectivity of the amygdala and surrounding structures reveals strong direct links with neural regions associated with modulating behavior on the basis of reward and punishment cues (e.g., basal ganglia, motor cortex, orbital frontal cortex; Davis & Whalen, 2001; Park & Judd, 2005) and for mobilizing fight or flight responses (e.g., via the hypothalamic–pituitary–adrenal axis; Feldman, Conforti, & Weidenfeld, 1995). By contrast, neocortical regions associated with semantic associations appear to have few, if any, direct connections to these systems. Rather, semantic associations are likely embedded in distributed networks in association cortex and thus may influence social cognition by biasing higher order information processing, such as when inferring the beliefs and intentions of another person (Amodio & Frith, 2006). Although systems for affect- and semantic-based associations typically function in concert, and thus tend to appear blended in outward verbal and behavioral responses, a consideration of their distinct operations is critical for understanding the behavioral effects of implicit stereotyping and evaluation.

Relationship Between Implicit Stereotyping and Implicit Evaluative Race Bias

On the basis of social psychological and neuroscientific theorizing, we proposed that implicit stereotyping and evaluation should represent independent constructs. Although past theorizing has pointed to this distinction (e.g., Greenwald & Banaji, 1995; Greenwald et al., 2002), few studies have explored it directly (cf. Dovidio et al., 1986; Kawakami et al., 1998; Rudman, Greenwald, & McGhee, 2001; Wittenbrink et al., 1997, 2001), and none has examined the respective implications of implicit stereotyping versus implicit evaluation for behavior. A limiting factor in this line of inquiry is that in previous research, independent assessments of implicit stereotyping and evaluative race bias have not been obtained, and thus it has not been possible to examine the conceptual relationship of implicit stereotyping and evaluative race bias and their potentially unique effects on behavior. Hence, the first major goal in the present work was to obtain independent measures of implicit stereotyping and evaluation that would permit a fair test of the independence hypothesis.

Differential Effects of Implicit Evaluative Race Bias and Stereotyping on Behavior

If implicit stereotyping and evaluation reflect independent cognitive and affective systems, then they may be uniquely associated with different types of discriminatory responses. Consistent with this hypothesis, Millar and Tesser (1986, 1989) argued that instrumental behaviors (e.g., forming judgments and goals) are driven primarily by cognitive processes, whereas consummatory behaviors (e.g., appetitive or aversive behaviors) are driven primarily by affective–evaluative processes. On the basis of this theorizing, Dovidio and his colleagues (1996, 2004; Esses & Dovidio, 2002; see also Stangor, Sullivan, & Ford, 1991) proposed that by considering the match between the affective versus cognitive nature of race-bias measures and forms of discriminatory outcomes, greater correspondences between assessments of race bias and behavior may be attained. In a meta-analysis focusing on explicit self-reports of stereotyping and prejudice, Dovidio et al. (2004) showed that affect-based measures of race bias tended to be predictive of basic approach/avoidance responses (e.g., nonverbal behaviors and affective responses) toward African Americans, whereas cognition-based measures of race bias tended to be predictive of the endorsement of stereotypes and support for policies that disadvantage African Americans.

Although Dovidio et al.'s meta-analysis focused on explicit measures of race bias, extant findings from the implicit race bias literature are generally consistent with these predictions (Ashburn-Nardo, Knowles, & Monteith, 2003; Dovidio et al.,

1997; Dovidio, Kawakami, & Gaertner, 2002; Fazio et al., 1995; McConnell & Leibold, 2001; Wilson, Lindsey, & Schooler, 2000). For example, Fazio et al. (1995) found that implicit evaluative bias was predictive of less friendly behavior toward a Black experimenter but was not associated with participants' views on the Rodney King verdict and ensuing riots. In other research, greater implicit evaluative bias was associated with more uncomfortable interactions (e.g., less eye contact, more blinking) with a Black confederate compared with those involving a White confederate (Dovidio et al., 1997, 2002) and more negative interactions with a Black experimenter on a host of indicators, including speech hesitations and errors and behavior judged to be abrupt, unfriendly, and uncomfortable (McConnell & Leibold, 2001). By contrast, researchers have not examined the unique effects of implicit stereotyping on behavior, although some previous findings bear on the topic. For example, Kawakami et al. (1998) found that higher levels of implicit stereotyping were predictive of the attribution of stereotypic traits to a larger proportion of African Americans (in addition to reporting more prejudiced attitudes). It is important to note that in previous research, the hypothesis that implicit stereotyping and implicit evaluation are uniquely predictive of alternative forms of race-biased behavior has not been directly tested. Hence, the second main goal of the present work was to test this hypothesis directly.

Overview of Present Research

In the present research, we examined the relationship between implicit stereotyping and implicit evaluative race bias and their respective effects on instrumental versus consummatory forms of race-biased behavior. Although stereotyping and evaluation processes typically operate in concert, it was necessary for us to obtain relatively pure measures of implicit stereotyping and evaluation to examine their unique effects on behavior. To this end, we designed separate implicit association tests (IATs) to assess implicit stereotyping and implicit evaluative race bias. The IAT was chosen because it has been shown to be reliable (Greenwald, Nosek, & Banaji, 2003), and it has been widely used in the implicit race bias literature (Devine, 2001). In Study 1, we examined the degree to which measures of implicit stereotyping and evaluative race bias were independent (i.e., uncorrelated). In Studies 2 and 3, we examined the unique effects of implicit stereotyping and evaluation on instrumental and consummatory forms of behavior.

Study 1

Method

Participants and Procedure

One hundred fifty-one European American introductory psychology students (82 women, 69 men) participated in exchange for extra course credit. After providing informed consent, participants received instructions on completing separate IAT measures of stereotyping and prejudice administered on a PC using Inquisit software (Millisecond Software, Seattle, WA). IAT order was counterbalanced across participants. After completing the measures, participants were debriefed, thanked, and dismissed.

Materials

Evaluative IAT. The IAT is a dual categorization task in which participants categorize words as pleasant or unpleasant and faces as either Black or White by pressing one of two keys on the computer keyboard. Stimuli consisted of pleasant and unpleasant words as used by Greenwald et al. (1998) and pictures of White and Black male faces displaying neutral expressions (Malpass, Lavigueur, & Weldon, 1973) as used by Devine, Plant, Amodio, Harmon-Jones, and Vance (2002; Study 3). Pleasant words included *honor*, *lucky*, *diamond*, *loyal*, *freedom*, *rainbow*, *love*, *honest*, *peace*, and *heaven*. Unpleasant words included *evil*, *cancer*, *sickness*, *disaster*, *poverty*, *vomit*, *bomb*, *rotten*, *abuse*, and *murder*.

The IAT procedure comprised five blocks of trials (Greenwald et al., 1998). Stimuli were presented individually in the center of the computer monitor in randomized order. In Block 1, participants viewed 10 Black and 10 White faces and categorized Black faces by pressing the left response key ("a" on the alphabetic keyboard) and White faces by pressing the right response key ("5" on numeric keypad). In Block 2, participants viewed 10 pleasant and 10 unpleasant words, categorizing unpleasant words with the left response key and pleasant words with the right response key. In Block 3, stimuli included White faces, Black faces, pleasant words, and unpleasant words, and response mappings were combined such that participants categorized Black faces and unpleasant words by pressing the left response key and White faces and pleasant words by pressing the right response key. This block consisted of 40 trials and was referred to as the *compatible* block (Greenwald et al., 1998), given that response pairings of White with good and Black with bad are compatible with Whites' tendency to prefer White faces over Black faces. In Block 4, participants viewed 10 Black and 10 White faces but this time categorized White faces with the left response key and Black faces with the right response key to counterbalance response mappings. In Block 5, categorizations were again combined such that participants categorized White faces and unpleasant words by pressing the left response key and Black faces and pleasant words by pressing the right response key. This block included 40 trials and was referred to as the *incompatible* block. Half of the participants completed the IAT as described above; half completed a version with reversed response mappings.

Stereotyping IAT. We designed a new IAT in which participants viewed two classes of words associated with the positive characteristics of intelligence and athleticism/rhythmicity, and categorized them as *mental* or *physical*, respectively, in addition to the Black versus White face categorizations. Intelligence and athleticism/rhythmicity are central to the African American stereotype, such that African Americans are stereotyped as more athletic/rhythmic and less intelligent than European Americans (Devine & Elliot, 1995). Because the mental and physical categories were relatively neutral, the categorization of words relating to athleticism/rhythmicity and intelligence as mental or physical did not involve evaluative judgments.[2]

Target word stimuli used in the stereotyping IAT were selected on the basis of pretesting.[3] Ten target words were selected for each category on the basis of category fit and stereotypicality. Mental words included *math*, *brainy*, *aptitude*, *educated*, *scientist*, *smart*, *college*, *genius*, *book*, and *read*. Physical words included *athletic*, *boxing*, *basketball*, *run*, *agile*, *dance*, *jump*, *rhythmic*, *track*, and *football*. The procedure for the stereotyping IAT was identical to that of the evaluative IAT, except that the pleasant and

[2] We developed additional IATs for other common African American stereotypes. Using the method by which Rudman, Greenwald, and McGhee (2001) measured implicit gender stereotyping, we pretested sets of target words related to poor (vs. wealthy), hostile (vs. friendly), and lazy (vs. motivated). In each case, however, the stereotype was strongly related to evaluation (e.g., poor is negative and wealthy is positive), and therefore these were not suitable for examining the independence of implicit evaluation and implicit stereotyping.

[3] Our lab group first generated separate lists of 22 words corresponding to the physical and mental categories. Sixty-one introductory psychology students then rated the fit of each word with its respective category and its degree of association with White and Black Americans on a scale of 1 (*not*

unpleasant target words and category labels were replaced with intelligence- and athletics-related target words and the *mental* and *physical* category labels. Hence, the compatible block included Black/physical and White/mental categorizations and the incompatible block included Black/mental and White/physical categorizations.

IAT scoring. Responses to the evaluative and stereotyping IATs were scored using the "improved algorithm," outlined by Greenwald et al. (2003, p. 214), which produced the *D* statistic.[4] However, because the IAT used in Study 1 consisted of the original five-block version (Greenwald et al., 1998), steps involving practice blocks were omitted. Following the algorithm, responses with latencies greater than 10,000 ms were removed. Separate means were computed for correct raw response latencies on compatible and incompatible blocks. Error responses within each block were replaced by the mean correct reaction time for that block, plus a 600-ms error penalty. *D* was quantified as the difference between incompatible and compatible mean reaction times divided by the pooled standard deviation of reaction times on compatible and incompatible blocks. Data from two participants were excluded because of outlying scores (Student's *t* scores differed significantly from mean, $p < .05$), and data from one participant were excluded because a high percentage of his responses (18%) on the stereotyping IAT were faster than 300 ms (Greenwald et al., 2003); results did not differ when outliers were included.

Results

Evidence for implicit bias was examined using one-sample *t* tests of *D* scores (effect size *r* is presented for each *t* value). All tests were two-tailed. Evaluative IAT scores were significantly greater than zero ($M = .51$, $SD = .42$), $t(147) = 14.60$, $p < .001$, $r = .77$, suggesting a negative evaluative association with Black faces relative to White faces, replicating past work. Stereotyping IAT scores were also significantly greater than zero ($M = .17$, $SD = .43$), $t(147) = 4.72$, $p < .001$, $r = .36$, such that participants exhibited a pattern of stereotypic trait associations with Black and White faces. No effects were found for sex or IAT order, Fs < 1. Next, we tested our primary hypothesis that levels of implicit stereotyping and implicit evaluation should be independent by examining their correlation. Participants' evaluative and stereotyping IAT scores were not significantly correlated, $r(147) = .06$, $p = .47$, supporting our hypothesis.

Discussion

The results of Study 1 showed that participants possessed significant levels of implicit evaluative and stereotyping biases but that their levels of each bias were uncorrelated, suggesting conceptual independence. It is noteworthy that although athleticism, rhythmicity, and (un)intelligence represent a subset of commonly observed African American stereotypes, they are among the most central to the stereotype. Indeed, these three attributes were the most frequently cited by participants instructed to freely list traits associated with African Americans (Devine & Elliot, 1995). Because our stereotyping IAT focused on the three most central traits of the African American stereotype, and given previous findings that the activation of a central stereotype typically activates the constellation of African American stereotypes (Devine, 1989; Lepore & Brown, 1997), it is likely that our measure of implicit stereotyping reflected associations with the general African American stereotype. Nevertheless, it would be important to show that stereotyping IAT scores were predictive of responses to an African American target, reflecting stereotypic content that reached beyond traits of (un)intelligence, athleticism, and rhythmicity.

Studies 2 and 3 were designed with two goals in mind: to replicate Study 1 findings and to test the hypothesis that implicit stereotyping and evaluation are uniquely predictive of different forms of race-biased behavioral outcomes. The behavioral effects of implicit stereotyping and evaluation in Studies 2 and 3 were examined using double-dissociation designs constructed to isolate unique effects of predictors on specific outcome variables. Here, we tested the hypothesis that implicit stereotyping would be associated with instrumental but not with consummatory forms of race-biased behavior, whereas implicit evaluative race bias would be associated with consummatory but not with instrumental forms of race-biased behavior.

Study 2

In Study 2, we examined the degree to which participants' levels of implicit stereotyping and evaluation influenced their impressions of an African American student. To measure instrumental forms of behavior, we assessed participants' use of stereotypes as they formed an impression of the African American student on the basis of the student's writing sample (Moreno & Bodenhausen, 2001). To measure basic approach/avoidance responses associated with consummatory behaviors, we examined participants' preference for the writer as a potential friend. We also collected participants' affective ratings of various ethnic groups, including African Americans, using a feelings thermometer measure. We

at all) to 9 (*extremely*). Pretest ratings of category fit for physical and mental target words exceeded the scale midpoint, ps $< .001$, indicating that target words were good exemplars of their respective categories, and fit scores for the mental and physical target words did not differ, $t(60) = .40$, $p = .69$, $r = .05$. Physical target words were rated as more stereotypical of Black people ($M = 7.68$, $SD = 0.98$) than of White people ($M = 5.19$, $SD = 1.26$), $t(60) = 13.12$, $p < .001$, $r = .86$, whereas mental target words were rated as more stereotypical of White people ($M = 7.23$, $SD = 1.42$) than of Black people ($M = 4.30$, $SD = 1.31$), $t(60) = 13.31$, $p < .001$, $r = .86$.

A separate sample of 39 participants rated the favorability of words associated with the mental/physical and pleasant/unpleasant IATs on a scale from 1 (*extremely unfavorable*) to 9 (*extremely favorable*). Pleasant words ($M = 7.93$, $SD = 0.54$) were rated much more favorably than unpleasant words ($M = 1.75$, $SD = 0.58$), $t(38) = 38.83$, $p < .001$, $r = .99$. Unexpectedly, mental words ($M = 7.06$, $SD = 0.78$) were rated as more favorable than were physical words ($M = 6.22$, $SD = .77$), $t(38) = 7.07$, $p < .001$, $r = .75$, although both mental and physical word lists were rated significantly above the neutral midpoint of the scale, ps $< .001$, and both were rated as more favorable than unpleasant words, ps $< .001$, and less favorable than pleasant words, ps $< .001$. Although mental words were rated more favorably than physical words, this difference was much smaller than the difference in ratings between pleasant and unpleasant words, $t(38) = 28.77$, $p < .001$, $r = .98$. We used covariate analyses in our hierarchical regressions to ensure that effects of stereotyping IAT scores were not driven by evaluative associations (and vice versa) because any shared variance was statistically controlled. If anything, the valence effect found among the stereotyping IAT words would enhance the relationship between implicit stereotyping and evaluation, thereby working against our hypotheses and rendering more conservative tests.

[4] IAT effects based on difference scores (e.g., Greenwald et al., 1998) replicated results reported for the *D* statistic in all studies. Analyses of difference scores are available from the authors.

hypothesized that implicit stereotyping but not implicit evaluation would be related to more stereotypic trait ratings of the African American student, whereas implicit evaluation but not stereotyping would relate to a greater desire to befriend the writer and more negative affective responses toward African Americans.

Method

Participants and Procedure

Thirty-six European American introductory psychology students (15 men, 21 women) participated in exchange for extra course credit. After providing consent, participants were told that the study consisted of two parts. The experimenter explained that the first part examined people's ability to form impressions of others on the basis of short writing samples. Participants were shown a set of 10 file folders containing different writing samples and were asked to choose one at random (although all folders contained identical materials). Participants were given the chosen folder, which contained the writer's demographic information, a copy of the essay, and a set of forms to record their ratings. The demographic information included the writer's name, age, sex, ethnicity, year in college, and hometown, indicating that he was a 19-year-old male African American sophomore from Milwaukee, WI. Participants transferred this demographic information onto the evaluation form; read the essay, which contained some grammatical and spelling errors; and then provided their ratings of the essay and the writer. As the second part of the study, participants completed the evaluative and stereotyping IATs, in counterbalanced order, and the feelings thermometer measure. The essay ratings, IATs, and the feelings thermometer measure were administered in this order to prioritize the more covert measures as a means of minimizing participants' suspicions. Lastly, participants were probed for suspicion regarding the cover story and hypotheses, debriefed, thanked, and dismissed. Five participants' data were excluded because their scores on one or more measures differed significantly from the mean ($p < .05$) in a Student's t distribution and were considered outliers. Although inclusion of outliers inflated standard errors and thus weakened effect sizes, it did not change the pattern of effects.

Materials

Evaluative and stereotyping IATs. The evaluative and stereotyping IATs consisted of the same stimuli described in Study 1 but were administered using DirectRT software (Empirisoft, New York) and included sets of 20 practice trials before the compatible and incompatible blocks. The *D* statistic was computed as in Study 1, with the additional incorporation of responses from the practice blocks (Greenwald et al., 2003).

Essay evaluation materials. The essay evaluation form included items for rating (a) the general quality and style of the essay (included to bolster the cover story), (b) the trait attributes of the writer, and (c) participants' liking of and perceived similarity with the writer. Trait ratings of the writer were made using a scale ranging from 1 (*not at all*) to 10 (*very much*) on a list of adjectives known to be highly associated with the Black stereotype (*lazy*, *dishonest*, *unintelligent*, and *trustworthy*; Devine & Elliot, 1995) intermixed with filler traits that were relatively neutral and not typically associated with the stereotype (*modest*, *assertive*, and *thoughtful*). Ratings were averaged to form separate indices of stereotypic ratings ($\alpha = .68$, with *trustworthy* reverse-scored) and neutral filler ratings ($\alpha = .53$). Liking ratings were made for five items (e.g., "The writer seems like the type of person I would like to get to know better"; "The writer and I have a lot of things in common") on a scale of 1 (*strongly disagree*) to 10 (*strongly agree;* mean ratings: $\alpha = .73$).

Feelings thermometer. The feelings thermometer questionnaire consisted of a scale along which a range of "degrees" were depicted, from 0° (*extremely unfavorable*) to 100° (*extremely favorable*), with 50° labeled *neither favorable nor unfavorable*. Ratings were provided for African Americans, European Americans, Asian Americans, and Latino Americans.

Results

IAT Effects

As in Study 1, participants exhibited significant levels of implicit evaluation ($M = 0.32$, $SD = 0.17$), $t(31) = 10.96$, $p < .001$, $r = .89$, and implicit stereotyping ($M = 0.29$, $SD = 0.23$), $t(31) = 7.24$, $p < .001$, $r = .79$, yet IAT scores were not significantly correlated, $r(30) = .16$, $p = .37$. No significant effects emerged for sex or IAT order, Fs < 2.04, ps $> .16$.

IAT Effects on Behavioral Responses

To test our main hypotheses regarding double dissociations of the stereotyping and evaluation IATs, we used hierarchical linear regressions. First, the *D* score for the IAT that was not hypothesized to predict the outcome was entered as a covariate in Step 1. In Step 2, *D* for the hypothesized predictor was added to the regression model. We could then obtain evidence for a double dissociation by examining the simultaneous effects of the two predictors in Step 2. The semipartial r (sr) is reported as an effect size estimate.

Stereotype Ratings

Evaluative IAT scores, entered in Step 1, did not predict stereotype ratings of the African American essay writer, $\beta = -.17$, $t(29) = -0.90$, $p = .37$, $sr = -.17$. However, higher stereotyping IAT scores were associated with more stereotypic ratings of the African American essay writer in Step 2, $\beta = .39$, $t(28) = 2.70$, $p = .03$, $sr = .39$, whereas the effect of evaluative IAT scores remained nonsignificant, $\beta = -.23$, $t(28) = -1.33$, $p = .20$, $sr = -.23$. Ratings of nonstereotypic traits were not associated with scores on the stereotyping IAT, $\beta = -.01$, $t(29) = -0.04$, $p = .97$, $sr = -.01$, or the evaluative IAT, $\beta = -.02$, $t(28) = -0.11$, $p = .92$, $sr = -.02$. Finally, when nonstereotypic ratings were included as a covariate in Step 1, stereotyping IAT scores continued to predict stereotypic ratings, $\beta = .39$, $t(27) = 2.73$, $p < .01$, $sr = .38$, whereas evaluative IAT scores did not, $\beta = -.18$, $t(28) = -1.13$, $p = .26$, $sr = -.18$.

Affective Responses

In analyses of preference for the writer, stereotyping IAT scores entered in Step 1 were not predictive of preferences, $\beta = .06$, $t(29) = 0.33$, $p = .75$, $sr = .06$. In Step 2, higher evaluative IAT scores were associated with less desire to befriend the essay writer, $\beta = -.32$, $t(28) = -1.79$, $p = .08$, $sr = -.32$, whereas the effect for stereotyping IAT scores remained nonsignificant, $\beta = .01$, $t(28) = .04$, $p = .97$, $sr = .01$, supporting our hypothesis. Participants' feelings thermometer ratings provided an additional index of consummatory response toward African Americans that could be used to corroborate the marginally significant effect on preference for the writer. Average thermometer ratings for Whites, Asians, and Latinos were entered in the first regression step as a baseline covariate, followed by stereotyping IAT scores in Step 2 and evaluative IAT scores in Step 3. The effect for baseline thermometer ratings was significant, $\beta = .90$, $t(29) = 11.40$, $p <$

.001, $sr = .90$, which reflected individual differences in scale usage, but the effect for stereotyping IAT scores was not significant, $\beta = -.10$, $t(28) = -1.25$, $p = .22$, $sr = -.10$. Notably, higher evaluative IAT scores were predictive of more negative feelings toward African Americans, $\beta = -.18$, $t(27) = -2.04$, $p = .05$, $sr = -.16$, consistent with effects for writer preference. Additional analyses examining IAT effects on thermometer ratings of Whites, Asians, and Latinos produced no significant effects.

Discussion

The results of Study 2 further supported the independence of implicit stereotyping and implicit evaluation, such that scores on the stereotyping and evaluative IATs were not significantly correlated. Furthermore, our regression analyses revealed the hypothesized double dissociation between implicit stereotyping and implicit evaluation effects. These results indicated that cognitive and affective forms of implicit race bias are uniquely associated with instrumental versus consummatory forms of race-biased behavior, respectively, and hence showed discriminant and predictive validity of the stereotyping and evaluative IATs. It is notable that although the stereotyping IAT focused on a subset of the African American stereotype (e.g., athleticism, rhythmicity, and lack of intelligence), it predicted a broader instantiation of the stereotype, including the descriptors of *lazy*, *dishonest*, and *(un)trustworthy*, consistent with research evidencing strong links between subcomponents of the stereotype (e.g., Devine, 1989).

Although Study 2 provided good support for our hypotheses using conventional social psychological measures, it may have been limited in some respects. For instance, the procedure of Study 2 did not provide a good model of how implicit race biases would predict a White person's responses in anticipation of a real-life interaction with an African American. A second potential limitation was that the predictor and outcome variables were collected in the same experimental session, precluding causal inference and raising the possibility that the outcome measures might have influenced IAT scores. These limitations were addressed in Study 3, in which participants completed measures of implicit stereotyping and evaluation several weeks before being recruited for a purportedly separate experiment in which they expected to interact with an African American participant.

Study 3

Study 3 comprised two sessions. In the first session, participants completed IAT measures of stereotyping and evaluative race bias. In the second, ostensibly unrelated session, participants were led to believe that they would interact with an African American partner on various tasks involving tests of academic (verbal and mathematic) and nonacademic (sports and popular culture) knowledge. Participants rated how well they thought that they and their partner would perform on each of these tasks (Ashburn-Nardo et al., 2003) as an index of instrumental behavior. To assess consummatory behavior, we measured the distance participants chose to sit from the partner's belongings in a row of chairs just prior to their interactions (Macrae, Bodenhausen, Milne, & Jetten, 1994). We hypothesized that implicit stereotyping but not evaluation would predict stereotype-consistent performance expectations, whereas implicit evaluation but not stereotyping would predict seating distance from the partner.

Method

Participants

In the first phase of this study, participants were 43 introductory psychology students, 23 of whom were successfully recruited later in the semester for what they believed was an unrelated study. Evaluative IAT data from 2 participants were missing because of a computer malfunction, leaving 21 participants (13 women, 8 men) with valid data from both sessions. IAT scores of participants who did versus those who did not return for Session 2 did not differ, $ps > .23$.

Procedure

Session 1. Participants completed stereotyping and evaluative IATs in one of two counterbalanced orders, and the IATs were scored to yield *D* scores, as in Study 2.

Session 2. Participants were told the study would involve pairs of participants. At the scheduled experiment time, the experimenter entered the waiting room and called out the names of the participant and the (imaginary) partner. The partner's name alternated between "Darnell Stewart" and "Tyrone Washington" to suggest African American ethnicity (Greenwald et al., 1998). Noting that the partner had not yet arrived, the experimenter escorted the participant to the experiment room to get started. After providing consent, the participant was told the following:

> We're studying peoples' ability to cooperate with another person on some tasks assessing different types of general knowledge. You and a partner are going to complete a set of tasks, and then your combined score on these tasks will be compared with other teams who are in this study. You should try your best on these tasks, because the teams with the top five combined scores will be entered into a drawing for $40.

Participants were then asked to rate their abilities in various subject areas, including their mathematic and verbal skills and their knowledge of sports and cultural trivia. The experimenter then left momentarily, purportedly to check for the arrival of the partner. After a few minutes, the experimenter returned to explain that the other participant had arrived and was filling out initial questionnaires in another room. The participant was then shown the one-page participant information form identical to that used in Study 2. The top half was already filled in by the partner so that the participant would see he was African American. The participant completed the bottom half of the form.

Next, the experimenter noted they were running behind schedule and gave the following explanation:

> To save time, I'm going to have you decide which tasks you'll do and which your partner will do. Then we'll all go to the main testing room. Remember, you want to choose tasks for yourself and your partner that will give you the best combined score, not just so that only you or he will do well. There are four different tests: one has questions from the math SAT, another has questions from the verbal SAT, and the other two have questions about sports and popular culture.

Participants indicated which tasks they chose for themselves and their partners and then rated their perceptions of how well they and their partners would perform on each of the tasks.[5] After leaving briefly to check up on the supposed partner, the experimenter explained that the participant and the partner would now meet together in another room to complete their tasks. The experimenter led the participant out of the experiment room and,

[5] Participants were rather egalitarian in their assignments of the academic tasks, with 20 of 21 participants assigning one of the SAT tasks to themselves and the other to the partner. This pattern restricted the variance of task assignments, and thus it was an insensitive measure of stereotype-consistent behavior.

explaining that the partner had left momentarily to use the bathroom, directed the participant to sit in one of a row of chairs to wait. Eight identical chairs were arranged in a line, equally spaced approximately 4 in. (10.16 cm) apart along the hallway. A coat and backpack putatively belonging to the partner were placed on the chair nearest to the experiment room doorway. After the participant chose a seat, the experimenter surreptitiously recorded the participant's seating position and then left momentarily to photocopy the participant's information sheet.

After returning, the experimenter explained that the session would have to end early and led the participant back into the experiment room. The experimenter then probed the participant for suspicion regarding the cover story and the connection between Sessions 1 and 2, provided a debriefing and full explanation of the procedures, and then thanked and dismissed the participant. Two participants expressed some suspicion but were unable to identify key aspects of the cover story, the connection between Sessions 1 and 2, or the hypotheses.

Materials

Participants rated how well they thought that they would perform on the tests of SAT mathematic and verbal skills, sports trivia, and popular culture on a scale ranging from 1 (*very poorly*) to 9 (*very well*). Ratings of expected enjoyment on each task were also made on a scale ranging from 1 (*not at all*) to 9 (*very much*). Next, participants rated their expectations of their partner's performance and enjoyment on the same tasks, using the same scales.

Results

IAT Scores

Participants exhibited significant levels of implicit evaluation ($M = 0.38$, $SD = 0.29$), $t(20) = 5.93$, $p < .001$, $r = .80$, and implicit stereotyping ($M = 0.15$, $SD = 0.18$), $t(20) = 3.70$, $p < .001$, $r = .64$. Evaluative and stereotyping IAT scores were uncorrelated, $r(19) = .02$, $p = .93$, replicating the findings of Studies 1 and 2.

Ratings of Partner Abilities and Enjoyment

An index was created to represent the extent to which the partner was expected to perform poorly on academic tasks but to excel on nonacademic tasks, relative to participants' own expected performance. Participants' self-expectation ratings on each task were subtracted from their partner-expectation ratings. These scores were standardized, with ratings of counter-stereotype skills (mathematic and verbal) reverse-scored, and averaged, such that higher scores represented more stereotype-consistent expectations of the partner's performance, relative to expectations of the self.

The hypothesized double-dissociation effects were tested using hierarchical regressions as in Study 2. In Step 1, evaluative IAT scores were not significantly associated with expectations of the partner's performance, $\beta = -.24$, $t(19) = -1.08$, $p = .29$, $sr = -.24$. However, in Step 2, higher stereotyping IAT scores significantly predicted more stereotype-consistent expectations for the partner's performance, $\beta = .47$, $t(18) = 2.32$, $p = .03$, $sr = .47$, whereas the effect of evaluative IAT scores remained nonsignificant, $\beta = -.25$, $t(28) = -1.24$, $p = .23$, $sr = -.25$.[6] When participant sex was included in Step 1 as a covariate, effects for sex, $\beta = -.18$, $t(18) = -0.76$, $p = .46$, $sr = -.17$, and evaluative IAT scores, $\beta = -.19$, $t(18) = -0.79$, $p = .44$, $sr = -.18$, were not significant, whereas the effect for stereotyping IAT scores remained significant, $\beta = .50$, $t(17) = 2.48$, $p = .02$, $sr = .49$. Similarly, ratings of expected partner enjoyment on more stereotype-consistent tasks were not associated with evaluative IAT scores in Step 1, $\beta = -.06$, $t(19) = -0.25$, $p = .81$, $sr = -.06$, but were significantly associated with stereotyping IAT scores in Step 2, $\beta = .44$, $t(18) = 2.07$, $p = .05$, $sr = .44$.

Seating Distance From Partner

On average, participants sat 1.7 ($SD = .78$) chairs away from the partner's belongings. Stereotyping IAT scores, included in Step 1, were not associated with seating distance, $\beta = -.09$, $t(19) = -0.37$, $p = .71$, $sr = -.09$. However, as revealed in Step 2, participants with higher evaluative IAT scores chose to sit further from the partner's belongings, $\beta = .44$, $t(18) = 2.10$, $p = .05$, $sr = .44$, whereas the effect of stereotyping IAT scores remained nonsignificant, $\beta = -.09$, $t(28) = -0.45$, $p = .66$, $sr = -.09$, supporting our hypothesis.

Discussion

The results of Study 3 corroborated and extended the findings of Study 2. Greater implicit stereotyping scores uniquely predicted more stereotype-consistent expectations for the partner's performance, whereas greater implicit evaluation scores uniquely predicted greater seating distance from the African American partner's belongings. These findings provided additional support for our double-dissociation hypothesis of implicit stereotyping versus evaluation, whereby implicit stereotyping is rooted in semantic processes and is uniquely predictive of discrimination associated with instrumental responses, whereas implicit evaluation is rooted in affective processes and is uniquely predictive of discrimination associated with consummatory responses.

The Study 3 findings allayed concerns over some potential limitations of Study 2. First, the differential effects of implicit stereotyping and evaluation of Study 2 were replicated in a more realistic, ecologically valid context. Second, the two-session procedure used in Study 3 alleviated concerns regarding the order in which measures were administered in Study 2. Moreover, because IAT scores collected in the initial session were predictive of behaviors weeks later, our results suggest the effects of implicit bias are stable over time.

General Discussion

The present research produced two major findings. First, results suggest that implicit stereotyping and evaluative race biases represent conceptually independent constructs. Despite exhibiting significant levels of bias on both implicit measures across studies, participants' scores on these two measures were not significantly correlated, consistent with evidence for independent mechanisms of basic cognitive and affective processes (Cacioppo et al., 1999; Squire & Zola, 1996; Zajonc, 1980). Second, the results showed that implicit stereotyping and implicit evaluation have unique effects on alternative forms of race-biased behavior. The results of Study 2 showed that implicit stereotyping but not evaluation was

[6] IAT scores were not associated with absolute ratings of expected performance for the self, $ps > .37$, or the partner, $ps > .30$, indicating that implicit stereotyping effects were observable only when partner ratings were anchored by participants' self-reference.

predictive of stereotype-consistent trait ratings of a Black student that were based on a short writing sample. In contrast, implicit evaluative race bias but not stereotyping was predictive of participants' belief that they would get along with the student as a friend. Study 3 extended these findings by focusing on participants' behavior as they prepared to interact with an African American partner. In this study, implicit stereotyping, but not evaluation, predicted stereotype-consistent expectations of how well the African American partner would perform on a series of tasks. On the other hand, implicit evaluative race bias, but not stereotyping, predicted how far participants chose to sit from the African American partner's belongings in a row of chairs. Although the samples used in Studies 2 and 3 were relatively small, the use of double-dissociation designs ensured that null effects were always interpreted in the context of a complementary significant effect, and therefore low statistical power cannot account for the pattern of results. Taken together, these findings support the overarching hypothesis that implicit stereotyping processes are predictive of instrumental forms of race-biased behavior, whereas implicit evaluative processes are predictive of consummatory forms of race-biased behavior.

Implications for Theory and Research on Implicit Race Bias

Clarifying the Construct of Implicit Race Bias

In recent years, social psychologists have grappled with the meaning of implicit race biases in an effort to understand what they represent, how they function, and what they may predict (cf. Devine, 2001; Fazio & Olson, 2003). Against a backdrop of mixed findings regarding the effects of implicit race bias on behavior (Blair, 2001), our theorizing and results suggest that significant effects of implicit race bias on behavior may be observed when their underlying affective versus cognitive processes are taken into consideration and are matched with classes of behavior associated with consummatory versus instrumental responses (cf. Ajzen & Fishbein, 1977). On the basis of neuroscientific research, implicit evaluation is supported by subcortical mechanisms and is most directly expressed in basic approach/withdrawal behaviors. By contrast, implicit stereotyping is supported by neocortical networks and is most directly expressed in biased cognitive processing. This analysis provides a theoretical basis for conceiving of implicit stereotyping and evaluation as independent constructs and suggests refined definitions of these constructs that are rooted in neural mechanisms of learning and memory. Furthermore, it suggests that the effects of implicit stereotyping versus evaluation are likely to be expressed to different degrees in different situations and on different assessments (Livingston & Brewer, 2002; Macrae, Bodenhausen, Milne, Thorn, & Castelli, 1997). Although findings to date regarding the effect of implicit race bias on behavior are notoriously mixed, many null effects reported in the literature may have resulted from a mismatch between forms of implicit bias with outcome measures of discrimination.

Implications for Theory

To date, theories of implicit race bias have not addressed the possibility that implicit forms of stereotyping and evaluation may arise from distinct underlying processes and may affect behavior via alternative routes of processing (cf. Greenwald et al., 2002). Granted, stereotypes and affective responses are typically congruent and work together to facilitate a coordinated response (e.g., racial discrimination). Nevertheless, the predictive utility of a theory depends on whether it can be used to discern underlying processes and their respective effects on behavior. Future models of implicit race bias will benefit from the conceptual distinction presented here in several ways. First, a consideration of alternative forms of implicit bias will enhance predictive validity by permitting more refined hypotheses for how different forms of implicit bias should affect behavior. Second, our analysis links implicit stereotyping and evaluative bias to physiological models of the brain and behavior, permitting integration with other theoretical approaches and suggesting appropriate physiological indicators for different forms of bias. Indeed, previous research has associated indices of amygdala activity with implicit evaluation (e.g., Amodio et al., 2003; Phelps et al., 2000). Although neural correlates of implicit racial stereotyping have not yet been determined, event-related potential research on stereotype-based expectancy violation is consistent with a neocortical (versus subcortical) substrate (e.g., Bartholow, Fabiana, Gratton, & Bettencourt, 2001).

If implicit stereotyping and evaluation arise from distinct neural substrates, as we proposed, it follows that they are learned and unlearned via different mechanisms. One may refine theories of implicit race bias malleability and change by considering the respective dynamics of classical (fear) conditioning versus semantic associative learning. For example, human and animal models of learning and memory suggest that implicit evaluations may be learned more quickly and unlearned more slowly than implicit stereotypes. They also suggest that claims that implicit prejudice can be extinguished following a single experimental manipulation may be implausible and that other interpretations should be considered (e.g., the manipulation inhibited the initial activation of bias or elicited preconscious forms of regulation).

Implications for Study Design

It follows from the theoretical implications listed above that future research will benefit from a careful selection of measures and response contexts when examining the effects of implicit bias on behavior. The results of the present work suggest that implicit evaluation corresponds most directly with consummatory responses involving basic behavioral approach/withdrawal and that these effects should be strongest when behaviors involve minimal controlled processing. By contrast, implicit stereotyping affects behavior by biasing cognitive processing and thus should be most evident on measures that involve a higher degree of cognitive processing, provided that participants are unaware of the potentially biasing effects.

Future Directions

Joint Effects of Implicit and Explicit Race Bias

Although we went to great lengths to distinguish effects of implicit stereotyping from those of implicit evaluation in the present work, these two forms of bias typically operate in concert. An important new theoretical issue concerns the interplay of implicit stereotyping and implicit evaluation: When and how do they operate in concert? For example, behaviors that combine elements of consummatory and

instrumental responses may be best predicted by the joint effects of implicit stereotyping and prejudice. Additionally, there are many situations in which explicit measures of prejudice and stereotyping may be better predictors of behavior. Finally, although levels of implicit stereotyping and evaluation were not correlated in our samples, these two forms of implicit bias may be more strongly correlated among some groups of individuals (e.g., highly biased individuals) than others. Future research is needed to explore how the full range of discriminatory behavior may be explained by complex interactions among implicit and explicit forms of prejudice and stereotyping for different groups of people.

Regulatory Mechanisms for Implicit Stereotyping Versus Implicit Evaluative Race Bias

Our findings raise new questions as to whether the behavioral effects of implicit stereotyping and evaluation may be regulated via different processes and whether either form of implicit bias is more difficult to regulate. It is likely that regulation occurs at several different levels. For example, the spreading activation of automatic stereotypes within a semantic network could be inhibited via lateral inhibition (Bodenhausen & Macrae, 1998). Alternatively, the effects of implicit stereotypes may be inhibited in behavioral channels, such that a stereotype-congruent response tendency is overridden by a deliberative unbiased response (Amodio et al., 2004). Implicit evaluation associated with amygdala activation may be inhibited by the countervailing activation of reward structures in the brain, or its effect on behavior may be overridden via controlled processes as a behavioral response is formed. The inhibition of implicit stereotyping and evaluation effects at the response-formation level likely rely on the same frontal cortical mechanisms of control (Amodio et al., 2004; Amodio, Kubota, Harmon-Jones, & Devine, 2006). On the other hand, the inhibition of stereotypes within a neocortical semantic network and evaluations within a subcortical affective network rely on different mechanisms, and thus the parameters of regulation may vary considerably. The present theoretical analysis highlights some previously unexplored complexities regarding mechanisms for regulating the effects of implicit race bias.

Conclusion

Affect and cognition represent two fundamental processes of the human mind, and the distinction between affective and cognitive processes is critical for the understanding of a wide range of psychological functions (Cacioppo et al., 1999). On the basis of past social psychological and neuroscientific theories, we showed that cognitive and affective components of implicit race bias are conceptually independent and are uniquely predictive of instrumental and consummatory forms of race-biased behaviors, respectively. The present work is also notable in that we took a social neuroscientific approach: We applied neurocognitive models of learning and memory to elucidate social psychological conceptions of implicit processes that had been poorly defined. Our findings suggest that greater conceptual clarity in implicit race bias research may be achieved by considering the differential effects of implicit stereotyping and evaluation when interpreting extant findings, developing new theories, and designing future research.

References

Ajzen, I., & Fishbein, M. (1977). Attitude–behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84,* 888–918.

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

Amodio, D. M., & Frith, C. D. (2006). *Meeting of minds: The role of the medial frontal cortex in social cognition. Nature Reviews Neuroscience, 1,* 268–277.

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eye blink responses and self-report. *Journal of Personality and Social Psychology, 84,* 738–753.

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science, 15,* 88–93.

Amodio, D. M., Kubota, J. T., Harmon-Jones, E., & Devine, P. G. (2006). *Alternative mechanisms for regulating racial responses according to internal vs. external cues. Social Cognitive and Affective Neuroscience, 1,* 26–36.

Ashburn-Nardo, L., Knowles, M. L., & Monteith, M. J. (2003). Black Americans' implicit racial associations and their implications for intergroup judgment. *Social Cognition, 21,* 61–87.

Bartholow, B. D., Fabiani, M., Gratton, G., & Bettencourt, B. A. (2001). A psychophysiological analysis of cognitive processing of and affective responses to social expectancy violations. *Psychological Science, 12,* 197–204.

Bechara, A., Damasio, H., & Damasio, A. R. (2003). Role of the amygdala in decision-making. *Annual Review of Neuroscience, 985,* 356–369.

Blair, I. (2001). Implicit stereotypes and prejudice. In G. Moskowitz (Ed.), *Cognitive social psychology: On the tenure and future of social cognition* (pp. 359–374). Mahwah, NJ: Erlbaum.

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 11, pp. 1–52). Mahwah, NJ: Erlbaum.

Bouton, M. E. (1994). Conditioning, remembering, and forgetting. *Journal of Experimental Psychology: Animal Behavior Processes, 20,* 219–231.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76,* 839–855.

Damasio, A. D. (1994). *Descartes' error: Emotion, reason, and the human brain.* New York: Avon.

Davis, M., & Whalen, P. J. (2001). The amygdala: Vigilance and emotion. *Molecular Psychiatry, 6,* 13–34.

Devine, P. G. (1989). Prejudice and stereotypes: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5–18.

Devine, P. G. (2001). Implicit prejudice and stereotyping: How automatic are they? Introduction to the special section. *Journal of Personality and Social Psychology, 81,* 757–759.

Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton Trilogy revisited. *Personality and Social Psychology Bulletin, 21,* 1139–1150.

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology, 82,* 835–848.

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

Dovidio, J. F., Brigham, J. C., Johnson, B. T., & Gaertner, S. L. (1996). Stereotyping, prejudice and discrimination: Another look. In C. N. McCrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 276–319). New York: Guilford.

Dovidio, J. F., Esses, V. M., Beach, K. R., & Gaertner, S. L. (2004). The role of affect in determining intergroup behavior: The case of willing-

ness to engage in intergroup affect. In D. M. Mackie & E. R. Smith (Eds.), *From prejudice to intergroup emotions: Differentiated reactions to social groups* (pp. 153–171). Philadelphia: Psychology Press.

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

Esses, V. M., & Dovidio, J. F. (2002). The role of emotions in determining willingness to engage in intergroup contact. *Personality and Social Psychology Bulletin, 28,* 1202–1214.

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

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

Feldman, S., Conforti, N., & Weidenfeld, J. (1995). Limbic pathways and hypothalamic neurotransmitters mediating adrenocortical responses to neural stimuli. *Neuroscience and Biobehavioral Reviews, 19,* 235–240.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, pp. 357–411). New York: McGraw-Hill.

Gabrieli, J. D. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology, 49,* 87–115.

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.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition. *Psychological Review, 102,* 4–27.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109,* 3–25.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197–216.

Hamilton, D. L. (1981). Stereotyping and intergroup behavior: Some thoughts on the cognitive approach. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 333–353). Hillsdale, NJ: Erlbaum.

Kawakami, K., Dion, K. L., & Dovidio, J. F. (1998). Racial prejudice and stereotype activation. *Personality and Social Psychology Bulletin, 24,* 407–416.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97,* 377–395.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23,* 155–184.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72,* 275–287.

Livingston, R. W., & Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology, 82,* 5–18.

Mackie, D. M., & Smith, E. R. (1998). Intergroup relations: Insights from a theoretically integrative approach. *Psychological Review, 105,* 499–529.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67,* 808–817.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., Thorn, T. M. J., & Castelli, L. (1997). On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Social Psychology, 33,* 471–489.

Malpass, R. S., Lavigueur, H., & Weldon, D. E. (1973). Verbal and visual training in face recognition. *Perception and Psychophysics, 14,* 285–292.

McConahay, J. B., & Hough, J. C. (1976). Symbolic racism. *Journal of Social Issues, 32,* 23–45.

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.

Millar, M. G., & Tesser, A. (1986). Effects of affective and cognitive focus on the attitude-behavior relation. *Journal of Personality and Social Psychology, 51,* 270–276.

Millar, M. G., & Tesser, A. (1989). The effects of affective-cognitive consistency and thought on the attitude-behavior relation. *Journal of Experimental Social Psychology, 25,* 189–202.

Moreno, K. N., & Bodenhausen, G. V. (2001). Intergroup affect and social judgment: Feelings as inadmissible information. *Group Processes and Intergroup Relations, 4,* 21–29.

Park, B. & Judd, C. M. (2005). Rethinking the link between categorization and prejudice within the social cognition perspective. *Personality and Social Psychology Review, 9,* 108–130.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, 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.

Rissman, J., Eliassen, J. C., & Blumstein, S. E. (2003). An event-related fMRI investigation of implicit semantic priming. *Journal of Cognitive Neuroscience, 15,* 1160–1175.

Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit stereotypes and prejudice. *Journal of Personality and Social Psychology, 81,* 856–868.

Rudman, L. A., Greenwald, A. G., & McGhee, D. E. (2001). Implicit self-concept and evaluative implicit gender stereotypes: Self and ingroup share desirable traits. *Personality and Social Psychology Bulletin, 27,* 1164–1178.

Spencer, S. J., Fein, S., Wolfe, C. T., Fong, C., & Dunn, M. A. (1998). Automatic activation of stereotypes: The role of self-image threat. *Personality and Social Psychology Bulletin, 24,* 1139–1152.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the USA, 93,* 13515–13522.

Stangor, C., Sullivan, L. A., & Ford, T. E. (1991). Affective and cognitive determinants of prejudice. *Social Cognition, 9,* 359–380.

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

Wittenbrink, B., Judd, C. M., & 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.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Evaluative versus conceptual judgments in automatic attitude activation. *Journal of Experimental Social Psychology, 37,* 244–252.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151–175.

Received April 29, 2004
Revision received January 9, 2006
Accepted January 27, 2006 ■