Self-Categorization With a Novel Mixed-Race Group Moderates Automatic Social and Racial Biases

Jay J. Van Bavel and William A. Cunningham

Pers Soc Psychol Bull 2009 35: 321 originally published online 19 December 2008
DOI: 10.1177/0146167208327743

Self-Categorization With a Novel Mixed-Race Group Moderates Automatic Social and Racial Biases

Jay J. Van Bavel
William A. Cunningham
The Ohio State University

People perceive and evaluate others according to social categories. Yet social perception is complicated by the fact that people have multiple social identities, and self-categorization with these identities shifts from one situation to another. Two experiments examined whether self-categorization with a novel mixed-race group would override automatic racial bias. Participants assigned to a mixed-race group had more positive automatic evaluations of Black ingroup than Black outgroup members. Comparing these evaluations to Black and White faces unaffiliated with either group indicated this preference was driven by ingroup bias rather than outgroup derogation. Moreover, both outgroup and unaffiliated faces elicited automatic racial bias (White > Black), suggesting that automatic evaluations are sensitive to both the current intergroup context (positive evaluations of novel ingroup members) and race (racial bias toward outgroup and unaffiliated faces). These experiments provide evidence that self-categorization can override automatic racial bias and that automatic evaluations shift between and within social contexts.

Keywords: *intergroup; social categorization; attitudes; social identity; implicit; automatic*

Prejudice has marked much of human history, from the feud between the Hatfields and McCoys to the pogroms of the Second World War. Over the past 50 years, psychologists have made considerable progress on this issue by exploring how functional features of social cognition, such as categorization, give rise to stereotypes and prejudice. Categorization involves sorting stimuli on the basis of similarity in a spontaneous and reflexive fashion and helps individuals to efficiently process an otherwise overwhelming amount of information and generalize existing knowledge to new stimuli (Bruner, 1957). To simplify the challenges of social living, people constantly and reflexively categorize others according to salient social categories such as race, gender, and age (Brewer, 1988; Fiske & Neuberg, 1990). Social categorization uses information associated with the current categorization rather than individual characteristics—compromising accuracy for efficiency and eliciting prejudiced perceptions, thoughts, and evaluations (Allport, 1954).[1] In two experiments, we assigned people to mixed-race groups to examine how social categorization can both elicit and erase intergroup biases in automatic evaluation. Specifically, we tested whether self-categorization with a novel group would elicit a preference for ingroup members and override automatic racial bias.

Development and Expression of Prejudice

In contemporary society, a host of social, cultural, and developmental factors sow the seeds for prejudice (Rudman, 2004b). Over a lifetime, exposure to stereotypes and prejudices generates deeply entrenched associations (Staats & Staats, 1958) that color the way people see, feel, and act toward others. These associations with race and other social categories are often so well learned that they are automatically activated when encountering members of these groups (Devine, 1989; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998) and influence behavior,

Authors' Note: Please direct correspondence to Jay J. Van Bavel, the Ohio State University, Department of Psychology, 1827 Neil Avenue, Columbus, OH 43210; e-mail: van-bave1.1@osu.edu.

PSPB, Vol. 35 No. 3, March 2009 321-335
DOI: 10.1177/0146167208327743

even among people who explicitly endorse egalitarian values (Gaertner & Dovidio, 1986; Pettigrew & Meertens, 1995). Thus, although automatic and controlled aspects of evaluation are often in agreement, they can be dissociated (Greenwald & Banaji, 1995; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Nosek, 2005), making automatic biases a virtually unavoidable aspect of intergroup perception. Moreover, people with the strongest automatic racial bias are more likely to engage in a wide variety of discriminatory behavior (Greenwald, Poehlman, Uhlmann, & Banaji, in press), including overt discrimination (Rudman & Ashmore, 2007).

The past half century of research on prejudice has painted a troubling picture: Ordinary cognitive processes trigger automatic racial biases despite egalitarian beliefs. Race, in particular, affects categorization within milliseconds (Ito & Urland, 2003) and appears to be highly salient and difficult to suppress (Park & Rothbart, 1982). This has lead several researchers to conclude that race is automatically encoded (Hewstone, Hantzi, & Johnston, 1991; Stangor, Lynch, Duan, & Glass, 1992) whereas other social categories (religion, occupation, etc.) may be easier to suppress. Furthermore, attempts to suppress racial bias often backfire, leading to mental exhaustion (Richeson & Shelton, 2003), increased use of stereotypes (Macrae, Bodenhausen, Milne, & Jetten, 1994), or worse—unfriendly interracial interactions (Norton, Sommers, Apfelbaum, Pura, & Ariely, 2006). It is therefore unsurprising that racial bias persists in real-world contexts, including hiring (Bertrand & Mullainathan, 2003) and mortgage lending (M. A. Turner et al., 2002).

Us Versus Them

People are remarkably adept at dividing up the world into *us* and *them,* even in the absence of any factors typically posited to account for intergroup bias, such as stereotypes, prior contact with ingroup or outgroup members, and competition over resources. In a series of classic studies, participants randomly assigned to groups on the basis of unimportant and arbitrary distinctions, such as whether they tend to overestimate or underestimate the number of dots on a screen, allocated more money to fellow ingroup than outgroup members (Tajfel, Billig, Bundy, & Flament, 1971). This minimal-group membership tends to elicit ingroup bias rather than outgroup derogation (see Brewer, 1979, for a review), even on automatic attitude measures (Ashburn-Nardo, Voils, & Monteith, 2001; Otten & Wentura, 1999; Perdue, Dovidio, Gurtman, & Tyler, 1990). Thus, automatic intergroup biases can emerge on the basis of simple intergroup distinctions in the absence of well-learned stereotypes and prejudices.

In social environments where there is less than complete racial integration, race or ethnicity may provide a visually salient cue to group membership (Cosmides, Tooby, & Kurzban, 2003; see also Sidanius & Pratto, 1999). When race covaries with the categorization of others into *us* versus *them,* race may become a particularly salient and stable basis for social perception and evaluation, combining well-learned associations with the human propensity for ingroup favoritism (Brewer, 1979; Brown, 1999). Alternatively, when race is unrelated to group membership, other arbitrary intergroup dimensions may drive social perception and evaluation (Kurzban, Tooby, & Cosmides, 2001; Sidanius & Pratto, 1999).

Several experiments by Kurzban et al. (2001) recently examined this possibility using a memory confusion paradigm (Taylor, Fiske, Etcoff, & Ruderman, 1978). Participants were asked to form impressions of eight individuals and saw a series of statements, each of which was paired with a photo of the individual who said it. Participants were then asked to recall which statements were made by each individual. The target individuals varied independently on race (half were White and half were Black) and group membership (half were members of one group and half were members of another). Consistent with previous research (Stangor et al., 1992), participants made more within-race than between-race errors during recall. That is, they were more likely to misattribute statements from one Black individual to another rather than to a White individual, indicating that they were using race to encode the statements. Participants also made more within-group than between-group errors, indicating that they were also using group membership to encode statements. However, when group membership was made visually salient by showing pictures of one group wearing yellow shirts and the other group wearing grey shirts, participants used group membership more than race to categorize individuals. This visually salient group distinction also decreased race-based encoding. Remarkably, brief exposure to salient intergroup coalitions elicited group-based categorization and made this a more potent social category than race, a category marked by years of exposure.

This research highlights the dynamic nature of social categorization and presents a method for reducing racial biases in evaluation. As we noted previously, people have multiple social identities, and self-categorization with any of these identities—however minimal—can shift from one situation to another (Tajfel, 1982). The current self-categorization, in turn, colors perceptions *and* evaluations of the self and others (J. C. Turner, Hogg, Oakes, Reicher, & Wetherell, 1987; J. C. Turner, Oakes, Haslam, & McGarty, 1994). Several related lines of research have manipulated self-categorization to reduce

racial bias. Research on the common ingroup identity model, for example, has shown that categorizing two separate groups (us and them) into an inclusive superordinate group (we) reduces self-reported intergroup bias (Gaertner, Rust, Dovidio, Bachman, & Anastasio, 1996). In similar vein, making multiple, cross-cutting social categories salient can lead people to perceive a shared social identity with outgroup members, reducing self-reported intergroup bias (Crisp & Hewstone, 2007). In fact, classic research found that assigning participants to mixed-gender groups on an arbitrary basis reduced gender discrimination (Deschamps & Doise, 1978, Study 2), showing that a presumably unimportant self-categorization can reduce self-reported intergroup bias toward a stronger, more important social categorization (see also Crisp et al., 2006). Although these studies were concerned with controlled aspects of evaluation, research suggests that automatic evaluations are highly sensitive to social context, including the salience of social categories (for reviews, see Blair, 2002; Fazio & Olson, 2003; Gawronski & Bodenhausen, 2006). We therefore predicted that assigning people to a mixed-race group should not only reduce race-based categorization and self-reported racial bias but also modify *automatic* racial bias by eliciting ingroup bias toward Black individuals who share a common social identity.

Research on multicategorizable targets has shown that social category salience moderates the automatic activation of underlying attitudes. For example, a series of experiments by Mitchell, Nosek, and Banaji (2003) found that categorizing Black athletes and White politicians according to *race* activates an automatic preference for *White* politicians, whereas categorizing the same targets according to *occupation* activates an automatic preference for Black *athletes*. However, conclusions about the nature of automatic evaluations to multicategorizable targets remain unclear for at least two reasons. First, research has shown that certain implicit attitude measures, such as the Implicit Association Test (Greenwald et al., 1998), evoke evaluations consistent with the category labels (see Olson & Fazio, 2003). Accordingly, automatic evaluations of multicategorizable targets in the research by Mitchell et al. may have been driven by attitudes toward the category labels rather than the spontaneous construal of multicategorizable targets. Second, although studies on multicategorizable targets suggest they are generally evaluated according to the most important or salient social category (Mullen, Migdal, & Hewstone, 2001; Urban & Miller, 1998), orthogonal social categories can influence intergroup evaluations in a variety of ways (for a review see Crisp & Hewstone, 2007). In the current research, participants' evaluations may stem from preexisting racial bias (White > Black), current salient self-categorization (ingroup > outgroup), the sum of these categories, or an interaction between race and self-categorization. However, if automatic evaluation of multicategorizable targets is highly sensitive to the current salient self-categorization, participants should have a preference for minimal ingroup members even when race is a competing, visually salient social category.

Overview and Objectives

We assigned participants to a novel mixed-race group and tested whether they would have positive automatic evaluations for Black and White ingroup members—overriding the standard pattern of automatic racial bias. Using a variant of the minimal group paradigm, participants memorized ingroup and outgroup members and then completed measures of their automatic and controlled evaluations of these multicategorizable targets. We needed participants to develop familiarity with faces before the evaluation tasks and ensure they identified with the novel group so we modified the classic minimal group paradigm by having participants learn the faces and told them the groups were competing. Using a similar paradigm, Kurzban et al. (2001) found that participants *categorized* targets according to group membership rather than race. However, Kurzban et al. examined the main effects of race and group membership on encoding and did not measure evaluations *or* test the interaction between group membership and race. It therefore remains possible that racial biases are still automatically activated when race is orthogonal to group membership or that group membership may affect evaluations of Black or White faces differently. The present research addresses these issues.

EXPERIMENT 1

Experiment 1 explores whether people spontaneously construe and automatically evaluate others according to a novel group membership when it is orthogonal to race. Participants were randomly assigned to a group or a control condition in which they merely saw the groups (replicating the paradigm from Kurzban et al., 2001). This allowed us to examine whether merely learning about the two mixed-race groups was sufficient to reduce the categorization and evaluation of multicategorizable targets according to race or whether identification with one of the groups was necessary for changing evaluations.

Method

Participants. Participants were 109 University of Toronto undergraduate students (84 females) who successfully completed the study for partial course credit.[2]

Procedure. Participants arrived at the lab and posed for a digital photograph. Participants were informed that they were in a study exploring how people learn about groups and they were randomly assigned to one of two groups (Lions or Tigers; $N = 72$) or a control condition ($N = 37$) in which they learned about the two groups without being assigned to one of them. Participants were also told that it was important for them to learn the members of both groups before moving to other phases of the study. Participants then completed two brief learning tasks (approximately 15 min) followed by evaluation measures. So that we could examine the spontaneous evaluations of these multicategorizable faces, participants completed automatic and controlled evaluation measures of the faces without explicit category labels (Olson & Fazio, 2003).

Learning. Participants learned about two mixed-race groups: there were 6 Black and 6 White males in each group. The 24 faces were randomly assigned to groups and were fully counterbalanced so that nothing in the appearance of the individuals allowed participants to visually sort them into groups (the experimental design logically guaranteed that participants were as likely to see individual faces as Lions or Tigers members). In the first learning task, participants spent 3 min memorizing the group membership of all 24 faces simultaneously: 12 members of Lions and 12 members of Tigers. In the second learning task, participants were presented with each of the 24 faces one at a time and indicated whether each face was a member of the Lions or Tigers. Participants in the experimental condition (i.e., assigned to a group) also saw and categorized a photo of their own face three times during the second learning task (randomly interspersed within the other 24 faces) to enhance identification with their group.[3] During the first set of trials, each face was presented with the group name (Lions or Tigers) on the computer monitor to enhance learning. During the second set of trials, the group name was removed so participants had to rely on their memory. Following each trial, feedback indicated whether the response was correct. After the learning tasks, participants completed measures of their automatic and controlled evaluations of the faces in counterbalanced order.

Automatic evaluation measure. So that we could measure automatic evaluations of the faces, participants completed a response-window priming task on a personal computer (Cunningham, Preacher, & Banaji, 2001; see Draine & Greenwald, 1998, for details). During this task participants were instructed to rapidly categorize each word as "good/liked" or "bad/disliked" (see Olson & Fazio, 2004). Participants were instructed to press 1 when a good word appeared and 2 when a bad word appeared. Following 24 practice trials, participants completed three critical blocks with 96 trials each. On each trial in these critical blocks a face from the learning task appeared for 150 ms (followed by a blank screen for 50 ms) before a positive (e.g., love) or negative (e.g., hatred) target word, which appeared for a 525-ms response window. Participants were instructed to ignore the faces. The dependent measure was the proportion of trials in which each participant correctly categorized the word as good or bad (response accuracy; $\alpha = .45$). To better estimate automatic evaluative processing, all responses that occurred after 600 ms were coded as incorrect because longer response times allow for more controlled processing (Neely, 1977) and we were interested in automatic evaluation. Following previous research (Draine & Greenwald, 1998), we assumed that faces with positive associations would increase accuracy to positive words and decrease accuracy to negative words.

Controlled evaluation measure. So that we could measure more controlled evaluations of the faces, participants rated each face. Participants were told that "people can often quickly determine who they like or dislike based on subtle facial features and expressions," and they were asked to rate each of the 24 faces on a 6-point liking scale (1 = *dislike* to 6 = *like*; $\alpha = .83$). Faces were presented in random order. We measured evaluations of individual faces to increase the correspondence between automatic and controlled measures (Fishbein & Ajzen, 1974). By assessing evaluations on individual faces with automatic and controlled measures we increased confidence that any dissociations between measures were due to differences in evaluative processing and not conceptual incongruities (e.g., responding to exemplars on the automatic measure and social categories on the controlled measure).

Results

Analysis strategy. Analyses of automatic or implicit attitude measures have tended to focus on mean-level differences in reaction time or accuracy. However, this approach reduces hundreds of trials to a single score for each participant, diminishing power and meaningful variance. To more accurately measure evaluation we used multilevel modeling (Goldstein, 1995). Multilevel modeling allows for the direct analysis of individual trials and helps overcome violations of independence that occur as a result of correlated trials within participants. When an assumption of independence is not satisfied, ignoring dependency among trials can lead to invalid statistical conclusions, namely, the underestimation of standard errors and the overestimation of the significance of

predictors (Cohen, Cohen, West, & Aiken, 2003). We therefore created multilevel models with trials nested within participants to provide more appropriate estimates of regression parameters. Multilevel modeling has been used successfully on automatic attitude measures in previous research (Dunham, Baron, & Banaji, 2006; Nezlek & Cunningham, 1998). Multilevel models were implemented in the SAS PROC MIXED procedure (see Singer, 1998). To assess automatic evaluations we conducted 2 (group: ingroup, outgroup) × 2 (race: Black, White) × 2 (target valence: positive, negative) repeated measures analysis on response accuracy for experimental participants, and a 2 (race: Black, White) × 2 (target valence: positive, negative) analysis for control participants.[4] To assess controlled evaluations we conducted a 2 (group: ingroup, outgroup) × 2 (race: Black, White) repeated measures analysis on liking for experimental participants and an analysis on race (Black, White) for control participants.

The effects of group and race on automatic evaluations. Following previous research (Cunningham et al., 2001), we expected control participants to show automatic racial bias on response-window priming. As anticipated, a Race × Valence interaction indicated that control participants had relatively more positive associations for White than Black faces, $F(1, 36) = 3.42$, $p = .07$, $d = .30$. Specifically, simple effects analyses confirmed that participants had relatively positive evaluations (i.e., higher accuracy to positive than negative words) for White faces, $t(36) = 2.37$, $p < .02$, $d = .34$, but were relatively neutral to Black faces, $t(36) = 0.08$, $p = .94$, $d = .01$ (see Figure 1, Panel A). Simply learning faces in a context in which race was orthogonal to group membership evoked automatic racial bias: a preference for White over Black faces.

The primary hypothesis in this experiment was that self-categorization with a novel mixed-race group would eliminate automatic racial bias. As seen in Figure 1, group membership did moderate automatic racial bias. Specifically, a Group × Race × Valence interaction, $F(1, 71) = 5.99$, $p < .02$, was characterized by automatic racial bias (a preference for White over Black faces) toward outgroup faces, $F(1, 71) = 6.40$, $p = .01$, $d = .32$, but not toward ingroup faces, $F(1, 71) = 0.87$, $p = .35$, $d = -.11$. Simple effects analyses confirmed that participants had more positive evaluations (i.e., higher accuracy to positive than negative words) of Black ingroup than Black outgroup faces, $F(1, 71) = 10.77$, $p < .01$, $d = .38$. These results demonstrate that membership in a mixed-race group can improve automatic racial evaluations. Simple effects analyses also revealed that participants had relatively positive automatic evaluations (i.e., higher accuracy to positive than negative words) for White ingroup faces, $t(71) = 2.24$, $p = .03$, $d = .28$; Black ingroup faces, $t(71) = 3.34$, $p < .001$, $d = .36$; and White outgroup faces, $t(71) = 1.96$, $p = .05$, $d = .24$, but did not have positive automatic evaluations for Black outgroup members, $t(71) = -0.98$, $p = .33$, $d = -.12$. Whereas White outgroup members were evaluated more positively than Black outgroup members—mirroring the pattern of results in the control condition—ingroup members were evaluated positively, regardless of race.

The effects of group and race on controlled evaluations. Despite evidence of automatic racial bias among control participants, we expected that controlled evaluations would be more egalitarian. As predicted, there were no strong racial biases on ratings of Black and White faces among both control, $t(36) = -0.69$, $p < .49$, $d = -.11$, and experimental, $F(1, 71) = 2.98$, $p < .09$, $d = -.20$, participants. If anything, experimental participants had a marginal preference for Black ($M = 3.49$, $SD = .70$) over White ($M = 3.34$, $SD = .63$) faces. These egalitarian racial evaluations were dissociated from the automatic racial bias. In contrast, we expected ingroup bias would be evident on both automatic and controlled measures. As predicted, experimental participants preferred novel ingroup ($M = 3.52$, $SD = .69$) to outgroup faces ($M = 3.31$, $SD = .68$), $F(1, 71) = 5.39$, $p = .02$, $d = .27$. These preferences were not qualified by a Race × Group interaction, $F(1, 71) = 0.00$, $p = .95$.

Automatic and controlled correlations. To examine the correlations between automatic and controlled evaluations among experimental participants, we computed mean automatic racial bias (White-positive × Black-negative – White-negative + Black-positive), controlled racial bias (White – Black), automatic ingroup bias (Ingroup-positive + outgroup-negative – ingroup-negative + outgroup-positive), and controlled ingroup bias (ingroup – outgroup) scores. There were modest correlations between automatic and controlled racial bias ($r = .17$, $p = .14$) and ingroup bias ($r = .15$, $p = .22$). Although these correlations were in the typical range of raw automatic-controlled correlations (Cunningham et al., 2001; Hofmann et al., 2005), they were not statistically significant.

Discussion

This experiment provides evidence that automatic evaluations are sensitive to a novel self-categorization within a complex intergroup context. Participants in the control condition, who were not members of either mixed-race group, had an automatic preference for White relative to Black faces (Cunningham et al., 2001), suggesting that they were categorizing faces according

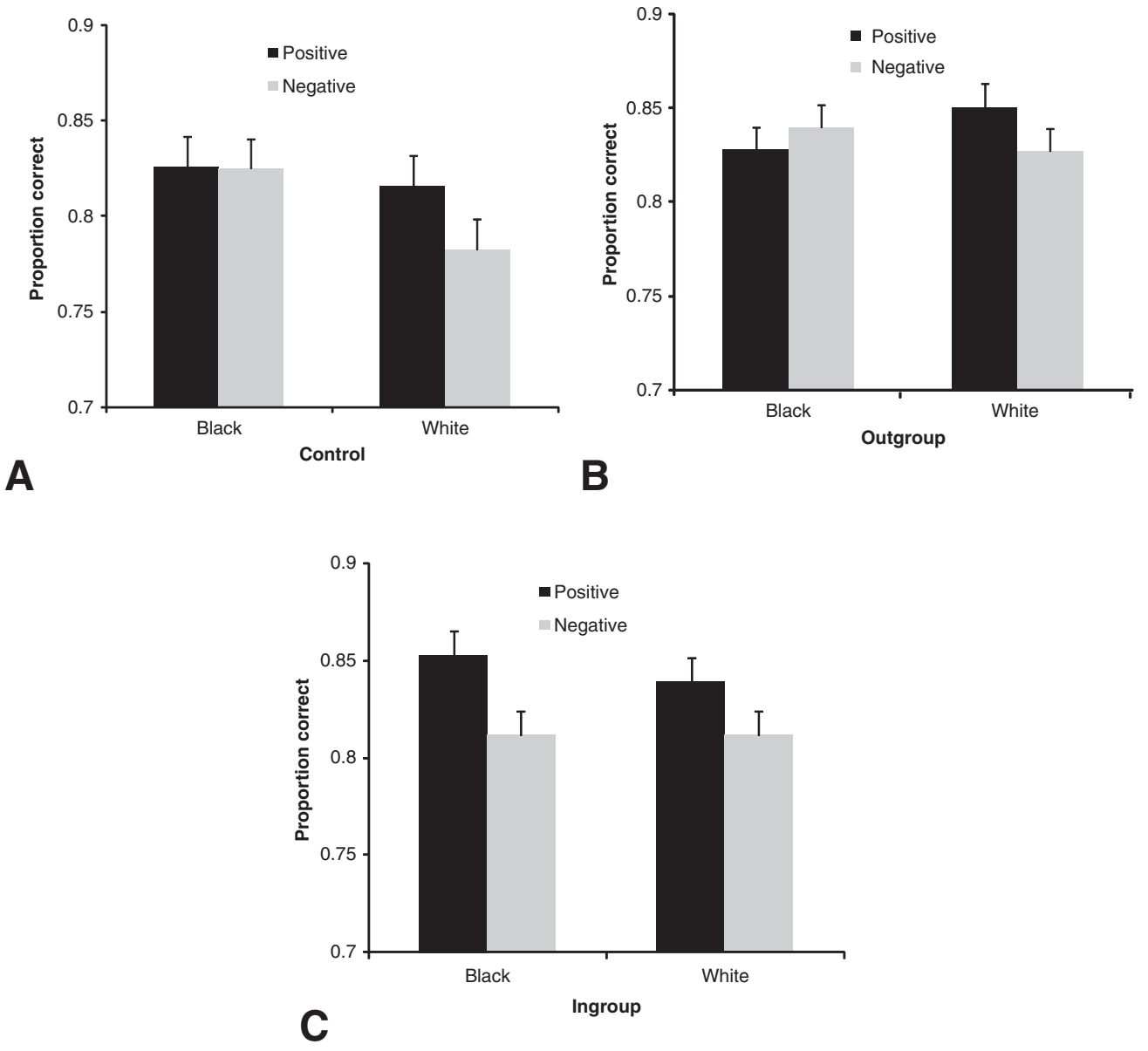

Figure 1 The effect of prime race (Black, White) and group membership (ingroup, outgroup) on response accuracy to positive and negative words (0.5 = chance responding) among control and experimental participants.

NOTE: Higher scores represent greater accuracy and lower scores represent less accuracy. (A) Control participants show the standard pattern of racial bias. (B) Experimental participants show the standard pattern of racial bias toward outgroup members. (C) Experimental participants evaluate ingroup members positively, regardless of race. Error bars show standard errors (Experiment 1).

to race—a visually salient social category. In contrast, participants in the experimental condition did not evaluate multicategorizable faces according to a single salient social category. Instead, their automatic evaluations were an interaction between self-categorization and race. They had positive evaluations for White and Black ingroup members, eliminating the standard pattern of automatic racial bias, whereas their automatic evaluations of outgroup members revealed racial bias, similar to the control condition. These results indicate that automatic evaluation was sensitive to self-categorization, shifting from evaluations based on race (control condition) to evaluations based on their current self-categorization (experimental condition).

Controlled evaluations were also sensitive to the current self-categorization: Participants assigned to a

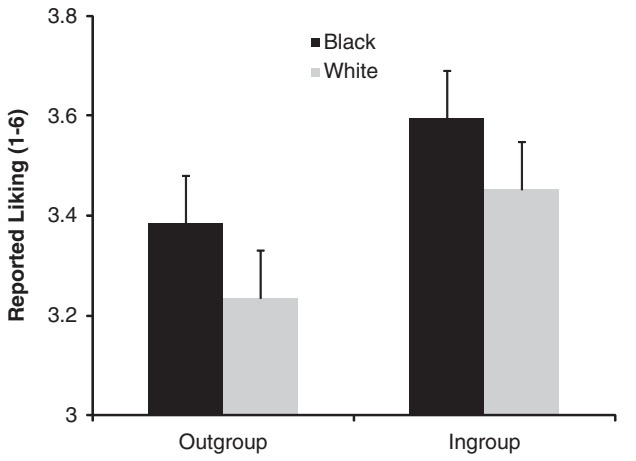

Figure 2 The effect of race (Black, White) and group membership (ingroup, outgroup) on self-reported liking on a 6-point scale (1 = *dislike* to 6 = *like*) among experimental participants.

NOTE: Participants show a preference for ingroup members relative to outgroup members, regardless of race. Error bars show standard errors (Experiment 1).

mixed-race group reported a preference for novel ingroup members, regardless of race. Replicating previous research (Devine, 1989), participants in the control condition revealed a dissociation between their automatic and controlled evaluations, showing an automatic preference for White over Black faces while reporting more egalitarian evaluations on controlled measures. We found this mean-level dissociation between automatic and controlled measures while having participants evaluate individual faces on both measures. In contrast, participants showed evidence of ingroup bias on automatic and controlled measures. These data suggest that racial bias can be modulated by more controlled processing or categorizing racial minorities as ingroup members.

This experiment offers evidence that automatic evaluations are sensitive to the current self-categorization, leading to automatic evaluations of ingroup members according to their group membership rather than race and generating more positive automatic evaluations of Black ingroup than Black outgroup members. These results are fully consistent with models that highlight the contextual sensitivity of self-categorization and the impact of these categorizations on social perception and evaluation (J. C. Turner et al., 1987), especially given the visual salience of race and extensive evidence that racial biases are automatically activated at the mere presentation of Black names or faces (Fazio et al., 1995).

EXPERIMENT 2

Experiment 2 sought to replicate our effects while examining how self-categorization comes to overshadow extant racial bias and whether the shift in evaluations was driven by more positive evaluations of Black ingroup members or more negative evaluations of Black outgroup members. A number of theoretical frameworks outline how self-categorization with a mixed-race group might reduce racial bias (Crisp & Hewstone, 2007; Gaertner & Dovidio, 2000; J. C. Turner et al., 1994). Although these frameworks share some common features, they offer competing predictions about the generality of self-categorization on intergroup evaluation. Specifically, if participants generate a superordinate ingroup identity (Gaertner & Dovidio, 2000) that includes all Black and White faces (except the current outgroup), they should have positive reactions to Black and White targets who are unaffiliated with either group. In contrast, if participants generate a more specific social identity (with the Tigers or Lions), they should show the standard pattern of automatic racial bias to Black and White targets who are unaffiliated with either group. To examine these alternative hypotheses, participants in Experiment 2 evaluated unaffiliated Black and White faces.

Experiment 2 was also designed to examine whether the automatic preferences for Black ingroup compared to Black outgroup members stem from ingroup bias, outgroup derogation, or both. Previous minimal group research has generally found ingroup bias—a preference for ingroup members (Brewer, 1979). For example, participants readily allocate more rewards to ingroup members but are reluctant to allocate more punishments to outgroup members (Mummendey et al., 1992), supporting the idea that evaluations of ingroup and outgroup members are not necessarily reciprocally related (Allport, 1954; Brewer, 1999). However, a meta-analysis on the evaluation of multicategorizable targets (Urban & Miller, 1998) suggested that more positive evaluations of crossed-category targets (e.g., Black ingroup members) are usually offset by more negative evaluations of double outgroup targets (e.g., Black outgroup members). To test these possibilities, we compared evaluations of ingroup and outgroup members with unaffiliated faces (i.e., Black and White faces that were not a member of either group). If self-categorization with a novel mixed-race group elicits ingroup bias, Black ingroup faces should be evaluated more positively than Black unaffiliated faces. However, if self-categorization elicits outgroup derogation, Black outgroup faces should be evaluated more negatively than Black unaffiliated faces.

Method

Participants. Participants were 126 undergraduate students from the Ohio State University who successfully completed the study for partial course credit. Two

participants were removed for using their iPod or cell phone during the experiment and 2 were removed for responding on less than 15% of trials during the priming task, leaving 122 (57 females) participants for analysis.

Procedure. The current experiment was similar to Experiment 1, with the primary difference that participants evaluated Black and White faces unaffiliated with either mixed-race group. Automatic ($\alpha = .75$) and controlled ($\alpha = .87$) evaluations of unaffiliated faces were contrasted with ingroup and outgroup members to identify the nature of preference for Black ingroup compared to Black outgroup members. Participants did not see unaffiliated faces during the learning phase of the experiment but saw them during both evaluation tasks. Participants were randomly assigned to one of two competing groups (Lions or Tigers) in the experimental condition ($N = 81$) or in a control condition ($N = 41$) in which they merely learned about the groups.

The procedure was different from Experiment 1 in three ways. First, participants memorized eight members of the Tigers and eight members of the Lions. Second, four Black and four White faces unaffiliated with either the Lions or the Tigers appeared for the first time during the evaluation tasks. Again, faces were randomly assigned to group and fully counterbalanced. Third, on the priming task half the participants pressed 1 when a good word appeared and 2 when a bad word appeared, and half the participants used the opposite word–number mappings.[5] Analyses were similar to Experiment 1 except that there were three groups (ingroup, outgroup, unaffiliated) of faces. To assess automatic evaluations we conducted a 3 (group: ingroup, outgroup, unaffiliated) × 2 (race: Black, White) × 2 (target valence: positive, negative) repeated measures analysis on response accuracy for experimental participants. To assess controlled evaluations we conducted a 3 (group: ingroup, outgroup, unaffiliated) × 2 (race: Black, White) repeated measures analysis on liking for experimental participants. Analyses contrasted unaffiliated faces with ingroup and outgroup members to identify the nature of preference for Black ingroup compared to Black outgroup members.

Results

The effects of group and race on automatic evaluations. The primary goals of this experiment were to replicate the results from Experiment 1 and examine how self-categorization comes to overshadow automatic racial bias. As we noted previously, if participants are generating a superordinate ingroup identity (Gaertner & Dovidio, 2000) that includes all Black and White faces (except the current outgroup), they should generalize their positive automatic evaluations to Black and White targets who are unaffiliated with either group. In contrast, if participants are generating a more specific social identity (with the Tigers or Lions), they should show the standard pattern of automatic racial bias to Black and White targets who are unaffiliated with either group. Replicating the results from Experiment 1, a Group × Race × Valence interaction, $F(2, 160) = 3.71$, $p < .03$, indicated that ingroup members were evaluated positively, regardless of race, $F(1, 80) = 0.85$, $p = .36$, $d = .08$, and more positively than unaffiliated faces or outgroup members (see Figure 3). More important, evaluations of outgroup and unaffiliated faces were characterized by automatic racial bias (a preference for White over Black faces), $F(1, 80) = 3.06$, $p < .08$, $d = .13$. These results suggest that participants were evaluating ingroup members according to their group membership (and not race) and outgroup and unaffiliated faces according to race rather than using a superordinate ingroup identity that generalized to novel targets.

We also examined the direction of automatic evaluation by comparing evaluations of ingroup and outgroup members with unaffiliated faces. As predicted, a marginal Group × Valence interaction, $F(1, 80) = 2.93$, $p = .09$, $d = .15$, indicated that participants had more positive automatic evaluations of ingroup members relative to unaffiliated faces, $F(1, 80) = 5.64$, $p < .02$, $d = .24$, but had similar evaluations for outgroup members and unaffiliated faces, $F(1, 80) = 2.78$, $p = .10$, $d = .19$; if anything, they had a marginal preference for outgroup members. Simple effects analyses confirmed that participants had more positive automatic evaluations (i.e., higher accuracy to positive than negative words) of Black ingroup than Black unaffiliated faces, $F(1, 80) = 12.74$, $p < .001$, $d = .39$, indicating that membership in a novel mixed-race group improved automatic evaluations of Black faces. These results are consistent with our hypothesis that ingroup members would be evaluated positively, regardless of race, and that evaluations of unaffiliated faces and outgroup members would not be significantly different, providing evidence of ingroup bias in the absence of outgroup derogation.

The effects of group and race on controlled evaluations. Following Experiment 1, we anticipated that participants would explicitly report racial egalitarianism toward all the faces (including unaffiliated) and a preference for novel ingroup faces relative to outgroup and unaffiliated faces. Participants reported a preference for ingroup members, $F(2, 160) = 6.19$, $p < .01$ (see Figure 4). Specifically, participants reported positive automatic evaluations of ingroup members ($M = 3.52$, $SD = .76$) relative to unaffiliated faces ($M = 3.29$, $SD = .70$), $t(80) = 3.78$, $p < .001$, $d = .44$, but nearly identical attitudes

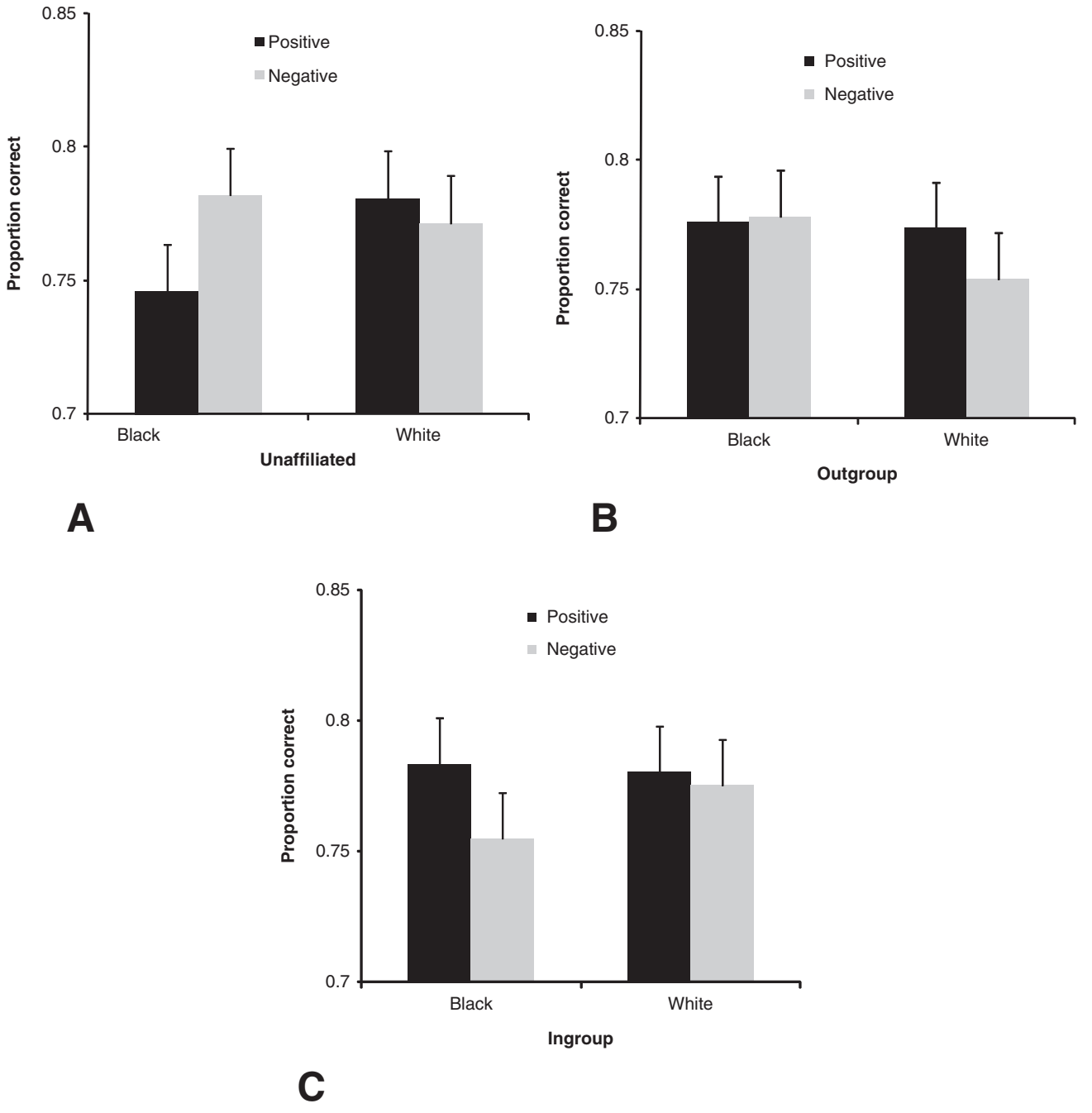

Figure 3 The effect of prime race (Black, White) and group membership (ingroup, outgroup, unaffiliated) on response accuracy to positive and negative words (0.5 = chance responding) among experimental participants.

NOTE: (A) Participants show the standard pattern of racial bias toward unaffiliated faces. (B) Participants show the standard pattern of racial bias toward outgroup members. (C) Participants evaluate ingroup members positively, regardless of race. Error bars show standard errors (Experiment 2).

toward outgroup members ($M = 3.28$, $SD = .79$) and unaffiliated faces, $t(80) = 0.14$, $p < .88$, $d = .02$. There was no main effect of race, $F(1, 80) = 1.72$, $p < .19$, $d = .15$, or Group × Race interaction, $F(2, 160) = 0.23$, $p = .79$. As expected, participants reported ingroup bias and egalitarian racial attitudes to all groups.

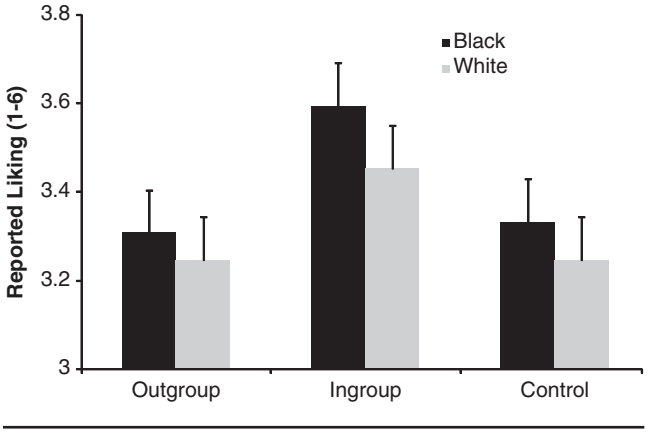

Figure 4 The effect of race (Black, White) and group membership (ingroup, outgroup, unaffiliated) on self-reported liking on a 6-point scale (1 = *dislike* to 6 = *like*) among experimental participants.

NOTE: Participants show a preference for ingroup members relative to outgroup members or unaffiliated members, regardless of race. Error bars show standard errors (Experiment 2).

Automatic and controlled correlations. Similar to Experiment 1, the correlations between automatic and controlled racial bias ($r = -.11$, $p = .32$) and ingroup bias ($r = .00$, $p = .97$) were not statistically significant.

Discussion

Replicating and extending the result from Experiment 1, participants had a preference for ingroup members, regardless of race, but continued to show automatic racial bias toward outgroup members and unaffiliated faces. In other words, positive evaluations of Black faces were limited to ingroup members and did not generalize to unaffiliated faces, suggesting that participants were not generating a superordinate ingroup identity but rather one specific to their novel group (Lions or Tigers). In addition, automatic and controlled evaluations were characterized by ingroup bias relative to unaffiliated faces (Brewer, 1979). Specifically, automatic evaluations of Black ingroup members were more positive than Black unaffiliated (and Black outgroup) faces and were not offset by more negative evaluations of Black outgroup members relative to Black unaffiliated faces. These results suggest that group membership increased positivity toward ingroup members without increasing negativity toward outgroup members.

GENERAL DISCUSSION

The current research examined the contextual sensitivity of automatic social evaluation in complex intergroup contexts and provided evidence that self-categorization with a novel mixed-race group can systematically alter social evaluation. In two experiments, people who were randomly assigned to a mixed-race group developed positive evaluations toward ingroup members, regardless of race. Specifically, group membership increased positivity toward Black ingroup members without increasing negativity toward Black outgroup members. These experiments suggest that automatic evaluations were highly sensitive to social context, reflecting the current salient intergroup alignment even when there were no visual differences between the ingroup and outgroup (a fact logically guaranteed by the experimental design); the groups had no history of contact or conflict; and there was an orthogonal, visually salient social category cue (i.e., race) with strong evaluative connotations. These results raise the possibility that mere categorization with a relatively unimportant group may be sufficient to override automatic evaluations of ingroup members according to race—a visually and social salient social category.

These experiments extended previous research showing that people were more likely to *categorize* group members according to group membership rather than race even if they were not a member of either group (Kurzban et al., 2001). The current research extended this earlier work by examining the effect of mixed-race groups on intergroup evaluations, analyzing the interaction between group membership and race, and actually *assigning* people to one of the mixed-race groups.[6] Participants in the control condition in Experiment 1 had an automatic preference for White compared to Black faces, suggesting that merely learning about mixed-race groups was insufficient to eliminate automatic racial biases. In contrast, participants who were assigned to a group showed evidence of automatic racial bias toward outgroup members and unaffiliated faces, but positive evaluations of Black and White ingroup members. Taken together, these results suggest that some aspect of self-categorization plays an important role in shifting more automatic aspects of social evaluation away from race.

Self-Categorization and Evaluation

These experiments join a growing body of research that has manipulated aspects of self-categorization to reduce intergroup bias. In the present research, participants assigned to a novel group evaluated both Black and White ingroup members positively, eliminating the standard pattern of automatic racial bias (Cunningham et al., 2001; Fazio et al., 1995). In contrast, participants in Experiment 1 who merely saw two mixed-race groups showed the standard pattern of automatic racial bias, highlighting the power of self-categorization and social identification to shape automatic evaluation. Taken together, these results

raise the possibility that participants were either generating a superordinate social identity that included all Black and White targets (except for outgroup members; Gaertner et al., 1996) or that the crossed-categorization of novel group membership and race leads to this reduction in racial bias (Crisp & Hewstone, 2007). If participants were generating a superordinate identity, they should show little or no racial bias toward any Black individual who is not an outgroup member. However, participants in Experiment 2 revealed automatic racial bias toward Black faces who were unaffiliated with the ingroup or outgroup. These data suggest that participants were not generating a superordinate social identity that extended beyond the current salient ingroup and raise the possibility that aspects of crossed-categorization may have accounted for the reduction in racial bias.

Crisp and Hewstone (2007) argue that there are two routes to reduced ingroup bias in crossed-category contexts: differentiation and decategorization. Differentiation involves the generation of a shared ingroup identity that brings the outgroup (i.e., Blacks) closer to the ingroup, resulting in reduced bias. Decategorization occurs in more complex intergroup settings and involves a shift toward more individuated processing (Fiske & Neuberg, 1990). Although the current research was not designed to investigate these potential mechanisms, participants assigned to a mixed-race group show greater activity in the fusiform gyrus—a brain region involved in individuation—when they see images of ingroup members, regardless of race (Van Bavel, Packer, & Cunningham, in press). Taken together, these experiments raise the possibility that a shared ingroup identity—however minimal—may lead individuals to feel positive about *and* individuate ingroup members. Future research should explore how these psychological processes mediate the effects of self-categorization on automatic racial bias.

The current research assumes that social perception and evaluation are closely linked. However, several recent studies have found an effect of crossed categories on categorization but not evaluation (e.g., Vescio, Judd, & Kwan, 2004), leading researchers to question the link between social categorization and intergroup bias (Park & Judd, 2005). Although it is likely that initial social perception does not have a 1:1 mapping with subsequent evaluative processing, the current research offers several insights for ongoing research on categorization and evaluation. First, the current research used measures of evaluation without explicit category labels, allowing multiple social categories (e.g., group and race) to drive evaluations. This minimized the extraneous influence of specific category labels on evaluation and allowed for spontaneous construals to impact evaluation. This approach revealed that participants spontaneously evaluated ingroup members according to their group membership but also that race was used during evaluation in a complex fashion. Therefore, the current crossed-category context did alter automatic evaluations. Second, the current research found that more controlled evaluations were not influenced by race. The dissociation may stem from additional component processes that alter evaluations to suit motivational concerns (Cunningham, Zelazo, Packer, & Van Bavel, 2007). In sum, this research suggests that links between categorization and evaluation will be closely linked when construals directly influence evaluation and number of component processes engaged in categorization and evaluation is similar, especially when assessing evaluations of social groups prone to motivated suppression or inhibition, like race (Crandall & Eshleman, 2003; Dunton & Fazio, 1997; Plant & Devine, 1998).

Experiment 2 examined the direction of biases to generate more precise inferences about the nature of intergroup preferences in complex social contexts, namely, whether automatic evaluations of Black faces were altered by ingroup bias and/or outgroup derogation. We presented unaffiliated Black and White faces during the evaluation tasks to contrast against ingroup and outgroup faces and found evidence of ingroup biases (i.e., positive evaluations of Black and White ingroup members) without outgroup derogation. In particular, ingroup faces were evaluated positively regardless of race, leading to more positive evaluations of Black ingroup faces, whereas outgroup members evoked evaluations similar to unaffiliated faces, the standard pattern of ingroup bias (Brewer, 1979). Although, a meta-analysis of crossed-categorization research (Urban & Miller, 1998) suggests that positive evaluations toward crossed-category targets (e.g., Black ingroup members) are usually offset by negative evaluations toward double outgroup targets (e.g., Black outgroup members), the pattern of ingroup bias absent outgroup derogation in the current research is consistent with the idea that ingroup and outgroup evaluations are not necessarily reciprocally related (Allport, 1954; Brewer, 1999).

Flexibility of Automatic Evaluation

Humans belong to many dynamic and overlapping social groups, and the importance of any given social category can shift between and within contexts. In such a complex and dynamic social world, a central challenge for adaptive human behavior is the flexible and appropriate categorization and evaluation of others. In the current research, participants assigned to a group had positive automatic and controlled evaluations of Black *and* White ingroup members, and these evaluative preferences were driven by ingroup bias and not outgroup derogation. In contrast, participants had more positive evaluations of White than Black outgroup

members and unaffiliated faces, suggesting that they were automatically construing and evaluating these targets according to their race. These results raise the possibility that automatic evaluations may be sensitive to different social categories for different faces, that is, the *group membership* of novel ingroup members and the *race* of novel outgroup members and unaffiliated faces. Moreover, these evaluations occurred while participants were randomly presented with a series of individual faces, providing evidence that automatic social perception and evaluation can rapidly and spontaneously shift between group membership and race on a case-by-case basis *within* a stable intergroup context.

These experiments suggest that self-categorization and automatic evaluations are flexible, leading to automatic evaluations that are both consistent with the current intergroup context and remain sensitive to the race of faces outside the ingroup. These findings advance the literature on intergroup evaluation in a number of ways. As noted earlier, much research on automatic evaluations has focused on relatively simple social categories or examined the automatic evaluations of multicategorizable targets with measures that use explicit category labels (e.g., Mitchell et al., 2003) that have been shown to elicit evaluations associated with the current label (Olson & Fazio, 2003). To examine more spontaneous evaluations of complex social stimuli, we used automatic and controlled measures of evaluation without explicit category labels. As anticipated, automatic evaluations of multiply-categorizable faces were complex, involving an interaction between a novel self-categorization and race. This interactive pattern of automatic evaluations suggests that people spontaneously and rapidly evaluate others according to a blend of contextually relevant self-categorizations and preexisting associations toward salient social categories.

Processing ingroup members according to their group membership, and outgroup members and unaffiliated faces according to their race, may occur for a number of reasons. One possibility is that outgroup members are perceived as less relevant or their group membership receives less attention, and they are thus processed according to more superficial, visually salient social category cues, such as race. A second possibility is that outgroup members are associated with relatively neutral evaluations, and thus race-based associations provide the basis for outgroup evaluations. These two possibilities imply a hierarchical model of intergroup perception and evaluation in which ingroup membership is processed first, and if individual faces do not belong to the ingroup, they are processed according to their race, whether because of an absence of attention or evaluative associations toward outgroup faces. A third possibility is that the current intergroup context may trigger positive automatic evaluations of *any* ingroup member, including novel ingroup members and White faces. In other words, any group associated with the self (whether racial or novel) elicits positive automatic evaluations. Although the current research does not speak to the specific mechanism behind this effect, the evidence does suggest that race is used to automatically evaluate outgroup members and unaffiliated faces.

Prejudice Reduction

Social cognitive approaches to prejudice reduction have focused heavily on the suppression of automatically activated evaluations and stereotypes (Monteith, Sherman, & Devine, 1998; Plant & Devine, 1998). However, suppression has proven a narrow and inefficient form of emotion regulation (Gross & Thompson, 2007) and can lead to increased stereotype accessibility (Macrae et al., 1994) or cognitive depletion (Richeson & Shelton, 2003) in intergroup contexts. The current research lends empirical support to alternative, antecedent-focused routes to prejudice reduction, namely, changing the ways that others are automatically construed. Changing automatic processing may be especially important because evaluation depends on information from early processes, and modest biases during the initial stages of perceptual and evaluation processing can have dramatic downstream effects (Cunningham et al., 2007). Although evaluation and behavior may feel controlled and deliberate, it is heavily informed by these initial automatic and nonconscious processes. The current experiments illustrate the power of a novel self-categorization to alter automatic components of social perception and evaluation and ultimately override ostensibly pervasive effects of race. Future research should examine whether this antecedent-focused approach to prejudice reduction bypasses the problems with suppression, including stereotype rebound and cognitive depletion.

Evidence that self-categorization with a novel mixed-race group elicits positive evaluations toward novel ingroup members without eliciting negative evaluations toward outgroup members suggests that mixed-race groups may provide a socially constructive mechanism for attenuating racial bias. However, automatic *evaluations* may be sensitive to the current salient self-categorization leading to a reduction in racial bias, whereas the underlying *attitudes* may remain relatively static (Cunningham et al., 2007). These properties allow for racial biases to unaffiliated faces but also for the return of racial bias toward ingroup members when group membership is no longer salient. It is worth noting that the social benefits of mixed-race group membership are offset by the caveat that self-categorization leads to ingroup biases in evaluation.

Although ingroup bias toward a novel group may seem like a fair trade-off for more pernicious racial biases, it is worth revisiting the effects of minimal group membership on more overt indices of intergroup discrimination (Tajfel et al., 1971). In many contexts, such as hiring or voting, any *differential* preference for one group over another may lead to a similar pattern of behavioral discrimination, whether it is driven by ingroup bias or outgroup derogation. Although self-categorization with a novel group may offer a simple and promising approach to reduce racial bias, it must be carefully weighed against the possibility of spawning new forms of intergroup bias.

Future Directions

The current research raises a number of interesting questions. As we noted previously, our experiments employed a modified minimal group paradigm (Tajfel et al., 1971): To enhance self-categorization, participants were told that the Lions and Tigers were in competition and they saw their own face appear during the learning task. It therefore remains an open question whether mere categorization is sufficient to override automatic racial bias or whether this categorization requires the additional salience of competition and self-images. Although the different results between the control and experimental conditions suggests that the self played an important role in shifting automatic evaluations from race to group membership (see Rudman, 2004a, for a more detailed discussion of the role of the self in automatic evaluations), it is unclear whether this pattern of ingroup bias stems from evaluative conditioning during the learning paradigm (where participants categorized their own photo and the faces of ingroup and outgroup members; Gawronski & Bodenhausen, 2006; Walther & Trasselli, 2003), self-anchoring (Otten & Epstude, 2006), or some other psychological process.

Future research should also explore whether the current pattern of results extend to other implicit attitude measures (Fazio et al., 1995; Payne, Cheng, Govorun, & Stewart, 2005) and social categories, including prejudices based on age, gender, nationality, and religion. Indeed, Kurzban et al. (2001) argue that the effects of certain social categories (i.e., age and gender) on social perception should be less malleable than race.

Conclusion

In an era of increasing globalization, social and economic harmony depends on the ability of people to cooperate with others from a variety of ethnic, geographic, and religious backgrounds. In such a complex and dynamic social world, a central challenge for adaptive human behavior is the flexible and appropriate categorization and evaluation of others. The current research illustrates that self-categorization with a social group—however minimal—can dramatically shift automatic social perception and evaluation, and override ostensibly pervasive racial biases. When participants were *not* a member of either mixed-race group, race seemed to serve as the default salient category among faces, and participants had an automatic preference for White relative to Black faces. However, when participants were assigned to a mixed-race group, they had an automatic preference for ingroup faces, regardless of race. These results suggest that automatic evaluations stem from a combination of preexisting associations toward social categories and flexible, self-categorization processes. Creating environments in which group membership is unrelated to existing social categories may help shift automatic evaluations away from more pervasive biases. This account is promising given the relationship between automatic racial biases and discrimination (Greenwald et al., in press) but must be weighed against the possibility of spawning new forms of intergroup bias.

NOTES

1. It was perhaps this observation that led novelist E. B. White to quip, "Prejudice is a great time saver. You can form opinions without having to get the facts."

2. The reported effects were not moderated by participants' gender.

3. Participants did not see their own face again during any other phase of the study.

4. We could not examine the 2 (condition: control, experimental) × 2 (group: ingroup, outgroup) × 2 (race: Black, White) × 2 (target valence: positive, negative) repeated measures analysis on response accuracy because control participants did not have ingroup or outgroup members.

5. Because participants in Experiment 1 always pressed 1 when a good word appeared and 2 when a bad word appeared, right-hand biases in information processing (Nisbett & Wilson, 1977) could have created the direction of these preferences. We counterbalanced response options during the priming task in Experiment 2.

6. Another difference is that there were no visual cues to distinguish the groups in the current research. Kurzban, Tooby, and Cosmides (2001) found reduced racial categorization when the two mixed-race groups were distinguished by uniform colors: One group wore gray jerseys, the other wore yellow jerseys. Therefore, reductions in racial bias may hinge on factors such as social category salience, and self-categorization may increase the salience of group membership in the absence of visual cues.

REFERENCES

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

Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *Journal of Personality and Social Psychology, 81,* 789-799.

Bertrand, M., & Mullainathan, S. (2003). *Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination* (Working Paper Series WP 03-22). Cambridge: Massachusetts Institute of Technology Department of Economics.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6*, 242-261.
Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86*, 307-324.
Brewer, M. B. (1988). A dual process model of impression formation. In T. Srull & R.Wyer (Eds.), *Advances in social cognition* (Vol. 1). Hillsdale, NJ: Lawrence Erlbaum.
Brewer, M. B. (1988). A dual process model of impression formation. In T. Srull & R.Wyer (Eds.), *Advances in Social Cognition* (Vol. 1, pp. 1-36): Earlbaum.
Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55*, 429-444.
Brown, D. E. (1991). *Human universals*. New York: McGraw-Hill.
Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64*, 123-152.
Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum.
Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences, 7*, 173-179.
Crandall, C. S., & Eshleman, A. (2003). A justification–suppression of the expression and experience of prejudice. *Psychological Bulletin, 129*, 414-446.
Crisp, R. J., & Hewstone, M. (2007). Multiple social categorization. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 39, pp. 163-254). Orlando, FL: Academic Press.
Crisp, R. J., Walsh, J., & Hewstone, M. (2006). Crossed Categorization in Common Ingroup Contexts. *Personality and Social Psychology Bulletin, 32*, 1204-1218.
Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12,* 163-170.
Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative reprocessing model: A multi-level framework for attitudes and evaluation. *Social Cognition, 25,* 736-760.
Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5-18.
Deschamps, J. C., & Doise, W. (1978). Crossed category memberships in intergroup relations. In H. Tajfel (Ed.), *Differentiation between social groups* (pp. 141-158). Cambridge: Cambridge University Press.
Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General, 127*, 286-303.
Dunham, Y., Baron, A. S., & Banaji, M. R. (2006). From American city to Japanese village: A cross-cultural investigation of implicit race attitudes. *Child Development, 77*, 1268-1281.
Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23*, 316-326.
Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013-1027.
Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology, 54,* 297-327.
Fishbein, M., & Ajzen, I. (1974). Attitudes toward objects as predictors of single and multiple behavioral criteria. *Psychological Review, 81,* 59-47.
Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1-74). New York: Academic Press.
Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism: Theory and researc*h (pp. 61-89). Orlando, FL: Academic Press.
Gaertner, S. L., & Dovidio, J. F. (2000). *Reducing intergroup bias: The common ingroup identity model*. Philadelphia, PA: Psychology Press.
Gaertner, S. L., Rust, M. C., Dovidio, J. F., Bachman, B. A., & Anastasio, P. A. (1996). The Contact Hypothesis: The role of a common ingroup identity on reducing intergroup bias among majority and minority group members. In J. L. Nye & A. M. Brower (Eds.), *What's social about social cognition?* (pp. 230-360). Newbury Park, CA: Sage.
Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692-731.
Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). London: Arnold.
Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.
Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74,* 1464-1480.
Greenwald, A. G., Poehlman, T. A., Uhlmann, E., & Banaji, M. R. (in press). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology.*
Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 3-24). New York: Guilford.
Hewstone, M., Hantzi, A., & Johnston, L. (1991). Social categorization and person memory: The pervasiveness of race as an organizing principle. *European Journal of Social Psychology, 21,* 517-528.
Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*, 1369-1385.
Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85*, 616-626.
Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences, 98,* 15387-15392.
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.
Mitchell, J. P., Nosek, B. N., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132*, 455-469.
Monteith, M. J., Sherman, J. W., & Devine, P. G. (1998). Suppression as a stereotype control strategy. *Personality and Social Psychology Review, 2*, 63-82.
Mullen, B., Migdal, M. J., & Hewstone, M. (2001). Crossed categorization versus simple categoization and intergroup evaluations: A meta-analysis. *European Journal of Social Psychology, 31*, 721-736.
Mummendey, A., Simon, B., Dietze, C., Grunert, M., Haeger, G., Kessler, S., et al. (1992). Categorization is not enough: Intergroup discrimination in negative outcome allocation. *Journal of Experimental Social Psychology, 28,* 125-144.
Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*, 226-254.
Nezlek, J. B., & Cunningham, W. A. (1998, October). *Studying individual differences in the activation of racial prejudice using the IAT and hierarchical linear modeling*. Paper presented at the meeting of the Society of Experimental Social Psychology, Lexington, KY.
Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.
Norton, M. I., Sommers, S. R., Apfelbaum, E. P., Pura, N., & Ariely, D. (2006). Color blindness and interracial interaction: Playing the political correctness game. *Psychological Science, 17*, 949-953.
Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134,* 565-584.

Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science, 14*, 636-639.

Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extrapersonal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology, 86*, 653-667.

Otten, S., & Epstude, K. (2006). Overlapping mental representations of self, ingroup, and outgroup: Unraveling self-stereotyping and self-anchoring. *Personality and Social Psychology Bulletin, 32*, 957-969.

Otten, S., & Wentura, D. (1999). About the impact of automaticity in the Minimal Group Paradigm: Evidence from affective priming tasks. *European Journal of Social Psychology, 29*, 1049-1071.

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.

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

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*, 277-293.

Perdue, C. W., Dovidio, J. F., Gurtman, M. B., & Tyler, R. B. (1990). Us and them: Social categorization and the process of intergroup bias. *Journal of Personality and Social Psychology, 59*, 475-486.

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

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75*, 811-832.

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

Rudman, L. A. (2004a). Social justice in our minds, homes, and society: The nature, causes, and consequences of implicit bias. *Social Justice Research, 17,* 129-142.

Rudman, L. A. (2004b). Sources of implicit attitudes. *Current Directions in Psychological Science, 13*(2), 80-83.

Rudman, L. A., & Ashmore, R. D. (2007). Discrimination and the Implicit Association Test. *Group Processes and Intergrroup Relations, 10*, 359-372.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 23*, 323-355.

Staats, A. M., & Staats, C. K. (1958). Attitudes established by classical conditioning. *Journal of Abnormal and Social Psychology, 57*, 37-40.

Stangor, C., Lynch, L., Duan, C., & Glass, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62,* 207-281.

Tajfel, H. (1982). *Social identity and intergroup behavior.* Cambridge, UK: Cambridge University Press.

Tajfel, H., Billig, M., Bundy, R., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1,* 149-178.

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36*, 778-793.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory.* Oxford, UK: Basil Blackwell.

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin, 20*, 454-463.

Turner, M. A., Freiberg, F., Godfrey, E. B., Herbig, C., Levy, D., & Smith, R. E. (2002). *All other things being equal: A paired testing study of mortgage lending institutions* (Report). Washington, DC: U.S. Department of Housing and Urban Development.

Urban, L. M., & Miller, N. (1998). A theoretical analysis of crossed categorization effects: A meta-analysis. *Journal of Personality and Social Psychology, 74*, 894-908.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of ingroup bias: An fMRI investigation. *Psychological Science, 19*, 1130-1138.

Vescio, T. K., Judd, C. M., & Kwan, V. S. Y. (2004). The crossed categorization hypothesis: Evidence of reductions in the strength of categorization but not intergroup categorization but not intergroup bias. *Journal of Experimental Social Psychology, 40,* 478-496.

Walther, E., & Trasselli, C. (2003). I like her, because I like myself: Self-evaluation as a source of interpersonal attitudes. *Experimental Psychology, 50*, 239-246.

Received April 26, 2008
Revision accepted September 10, 2008