Research Article

Why Are Some Individuals Not Racially Biased?

Susceptibility to Affective Conditioning Predicts Nonprejudice Toward Blacks

Robert W. Livingston[1] and Brian B. Drwecki[2]

[1]Northwestern University and [2]University of Wisconsin–Madison

ABSTRACT—*Two studies investigated the psychological underpinnings of racial nonbias, defined as extremely low or null bias on measures of implicit and explicit racial attitudes. In Study 1, racially nonbiased Whites showed differential susceptibility to affective conditioning compared with subjects with greater bias. A significant two-way interaction emerged, indicating that nonbiased individuals were significantly less likely than other individuals to acquire negative affective associations to neutral stimuli in a classical conditioning paradigm, but were more likely than other individuals to acquire positive affective associations to neutral stimuli. This pattern of findings was replicated in Study 2, in which the identification of nonbiased Whites was facilitated by their nomination by an African American acquaintance. Implications for bias formation and prejudice reduction are discussed.*

For most of its history, social psychology has focused on the evolutionary, psychological, economic, and social determinants of prejudice and intergroup discrimination (for reviews, see Brewer & Brown, 1998; Fiske, 1998; and Sidanius & Pratto, 1999). Investigations of whether and why some individuals do not exhibit negativity toward out-groups have been much more rare. In his classic treatise *The Nature of Prejudice*, Allport (1954) wrote that "it is the pathology of bigotry and not the wholesome state of tolerance that, as a rule, interests social scientists" (pp. 425–426). More than 50 years later, only a handful of studies have directly examined the topic of racial tolerance, and most of these studies have focused on higher-order values, such as egalitarianism (Monteith & Walters, 1998), universal orientation (Phillips & Ziller, 1997), or motivations to avoid prejudice (Plant & Devine, 1998), as explanations for tolerance. No previous study has explored the basic psychological processes underlying racial nonbias. In the present article, we adopt a positive-psychology approach to understanding the etiology of prejudice by examining the cognitive processes that enable some individuals to avoid racial bias on all levels.

We focused our investigation on nonbiased individuals because we believe that this population is psychologically unique. Although prejudice may be morally reprehensible by contemporary Western standards, scientific research suggests that intergroup bias is the rule rather than the exception. Several lines of research confirm that (a) discriminatory bias is ubiquitous across history, regions, and cultures (e.g., Sidanius & Pratto, 1999); (b) (implicit) discriminatory bias is observable in most individuals (e.g., Nosek, Banaji, & Greenwald, 2002); (c) humans are cognitively predisposed to intergroup bias (e.g., Macrae, Milne, & Bodenhausen, 1994; Tajfel & Turner, 1986); and (d) in-group bias is logical and functional insofar as it serves to maintain or enhance social, economic, and political resources (Sidanius & Pratto, 1999). Moreover, negative images of Blacks in American culture and the media render anti-Black bias even more probable (Devine, 1989). In short, the numerous propensities, incentives, and conditions for racial bias raise the question of whether there are individuals who are *not* biased, and if so, *why* they are not biased.

To be sure, the pervasiveness of racial bias depends on how it is conceptualized and measured. When conscious or explicit attitudes are measured, very few Whites show evidence of anti-Black bias (Fazio, Jackson, Dunton, & Williams, 1995). However, when automatic or implicit attitudes are measured, the inverse is true (Devine, 1989; Fazio et al., 1995), with more than 80% of Whites showing significant evidence of anti-Black bias

Address correspondence to Robert W. Livingston, Kellogg School of Management, Northwestern University, 2001 Sheridan Rd., Evanston, IL 60208, e-mail: rwlivingston@kellogg.northwestern.edu.

(Nosek et al., 2002). Devine (1989) has argued that individuals who consciously disavow prejudice are nonprejudiced, even if their nonconscious or automatic responses to Blacks are negatively biased. In contrast, Fazio et al. (1995) defined "truly" nonprejudiced individuals as those who do not show automatically activated negative associations to African Americans, and further argued that self-report measures reflect socially desirable standards of response, rather than actual racial attitudes.

In the present study, we operationalized nonbias as very low or null bias on both implicit and explicit measures of racial attitudes.[1] Recent research has indicated that both types of measures are valid, but that they tap qualitatively distinct attitudes and predict different discriminatory outcomes (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997). We took a conservative empirical approach by comparing the cognitive processes of a nonbiased sample with those of all other subjects (i.e., those who did not fall into the nonbiased category), rather than with those of a selected sample of subjects high in bias. Although there are likely to be a number of paths to racial nonbias, involving both social and biological factors, in the present study we focused on cognitive predictors, namely, susceptibility to affective conditioning, category-based processing, and inhibitory ability.

Because one source of attitude formation is associative learning (Zanna, Kiesler, & Pilkonis, 1970), it is possible that mechanisms underlying the formation of general affective associations also govern the formation of racial biases. A long history of attitude research has implicated classical conditioning as a mechanism by which attitudes can be acquired and maintained (Cacioppo, Marshall-Goodell, Tassinary, & Petty, 1992; Staats & Staats, 1958; Zanna et al., 1970), and more recent research has shown that conditioning processes are closely tied to the genesis of racial attitudes (Olson & Fazio, 2001, 2006; Olsson, Ebert, Banaji, & Phelps, 2005). One possibility is that individuals who are susceptible to the acquisition of negative affective associations in general are more prone than other individuals to develop racial biases toward stigmatized groups. Although a couple of studies have examined individual differences in affective orientation (Ito & Cacioppo, 2005; Zinbarg & Mohlman, 1998), no previous research has investigated whether individual differences in affective orientation predict racial attitudes.

A second possibility is that nonbiased individuals are more "color-blind" (i.e., less predisposed to think of others categorically) than ordinary individuals. Some researchers have argued that prejudice is the product of humans' reliance on categories as devices for simplifying an exceedingly complex world (Allport, 1954; Macrae et al., 1994). The social cost of perceiving individuals as instances of social categories is the activation of stereotypes and prejudice, particularly at the nonconscious or

automatic level (Devine, 1989). The tendency to categorize individuals into in-groups and out-groups, and to favor members of the in-group, is well documented (see Brewer & Brown, 1998, for a review). Indeed, in-group bias can even be obtained for *minimal* groups, or meaningless categories that are artificially created and randomly assigned (Tajfel & Turner, 1986).

Finally, it is possible that nonbiased individuals actually hold the same biases as ordinary individuals, but are better self-regulators. Previous research has shown that greater inhibitory ability is related to lower levels of bias toward African Americans, whether inhibition is measured via performance on cognitively demanding tasks (Richeson & Trawalter, 2005) or via age-related inhibitory differences (von Hippel, Silver, & Lynch, 2000), or manipulated by consumption of alcohol (Bartholow, Dickter, & Sestir, 2006). Furthermore, this relationship has been found for both implicit (Bartholow et al., 2006; Richeson & Trawalter, 2005) and explicit (von Hippel et al., 2000) measures of racial bias. Thus, greater inhibitory ability among nonbiased individuals may be evidence that these individuals are simply more able than others to regulate their racial biases.

Study 1 tested the extent to which each of these mechanisms predicted Whites' racial nonbias toward African Americans.

STUDY 1

Method

Subjects

The subjects were White introductory-psychology students at the University of Wisconsin who participated in exchange for course credit. The criteria for designating subjects as nonbiased were based on quartile scores of the subject population (i.e., undergraduates enrolled in introductory psychology courses) on a measure of explicit bias, the Attitudes Toward Blacks scale (ATB; Brigham, 1993), and on a measure of implicit bias, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). Our criteria for nonbias were an IAT score in the lowest quartile for students previously tested in our laboratory (i.e., an approximate difference score of less than 150 ms) and an ATB score in the upper quartile. Because these two measures are not highly correlated, roughly 7% of the student population would meet the combined criteria for nonbias by default. Therefore, we developed strategies (both in this study and in Study 2) to "oversample," or increase the percentage of individuals in the nonbiased category. All subjects whose scores were outside the designated quartile range on one or both of these measures of bias were designated as ordinary.

Because past research has demonstrated that scores on the Internal and External Motivation to Respond Without Prejudice scales (IMS and EMS, respectively; Plant & Devine, 1998) interact, such that high-IMS/low-EMS individuals show low levels of implicit and explicit racial bias (see Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002, for a detailed discussion), we

---

[1]Because it is difficult to empirically confirm the absolute absence of racial bias (or any psychological construct), we use "nonbias" more as a descriptive term than as a literal assessment of bias.

preselected 31 individuals who scored in the upper tercile of the former scale and the lower tercile of the latter scale from a pool of subjects who had completed these scales in a mass pretesting session 2 to 8 weeks earlier. This preselection process yielded a larger percentage of nonbiased individuals (i.e., more than 40%) than would be expected in a quasi-random recruitment process. Other subjects ($n = 79$) were recruited via a sign-up sheet for an experiment on "social judgment." Although the vast majority of these students fell into the ordinary category, a small percentage met the criteria for nonbias.

The final sample consisted of 110 students, 18 nonbiased and 92 ordinary.

Measures

All measures were administered on computers via MediaLab or DirectRT (Jarvis, 2004).

Measures for Categorizing Subjects. The IAT is a computerized measure of implicit bias that assesses reaction times to categorize words. Responses to two categories are made with the same key, and responses to two other categories are made with a second key; the pairing of categories is then switched (see Greenwald et al., 1998, for a full description). The critical measure in this study was the difference between reaction times when stereotypically Black names, such as *Jamel*, were paired with pleasant words, such as *peace* (and stereotypically White names, such as *Josh*, were paired with unpleasant words, such as *pain*) and reaction times when the pairings were reversed. Following the IAT, subjects completed the ATB scale ($\alpha = .85$). This 20-item inventory assesses subjects' agreement with statements such as, "It would not bother me if my new roommate was Black."

Susceptibility to Affective Conditioning. Subjects completed a susceptibility-to-affective-conditioning task (i.e., the SAFCON task) that we created to assess the facility with which individuals acquire positive and negative affective associations to neutral stimuli. All positive and negative images were selected from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2005). Nine Chinese ideographs were paired with 3 positive images, 3 negative images, and 3 pictures of neutral, gray squares in a random and balanced fashion to form a set (see Table 1). Three sets were created to ensure that each ideograph was paired with positive, negative, and neutral pictures across subjects. Each subject was randomly assigned to one of the three sets.

The display sequence during the task was as follows: presentation of a Chinese ideograph for 150 ms; presentation of a blank display for a 125-ms interstimulus interval; presentation of a positive, negative, or neutral prime (i.e., photo pairing) for 200 ms; and presentation of a blank display for a 200-ms interval before the presentation of the next ideograph. A given ideograph-prime pairing occurred five times in a row and constituted

TABLE 1

Pairings of the Chinese Ideographs With Positive, Negative, and Neutral Primes

Chinese ideograph	Set 1	Set 2	Set 3
往	−	+	N
松	−	+	N
刻	−	+	N
纪	+	N	−
既	+	N	−
含	+	N	−
波	N	−	+
承	N	−	+
射	N	−	+

Note. Each ideograph was paired with one positive (+), one negative (−), and one neutral (N) prime across the three conditioning sets.

one series. Subjects were presented with each series twice, in random order. Thus, each subject was shown a total of 90 ideograph-prime pairings (30 positive, 30 negative, and 30 neutral). After the presentation of all 18 series, subjects rated their liking of the 9 Chinese ideographs on a Likert-type scale from 1, *strongly dislike*, to 7, *strongly like*.

Categorization. Subjects completed a version of the who-said-what paradigm (Taylor, Fiske, Etcoff, & Ruderman, 1978) in which they viewed a simulated conversation (with written phrases presented on a computer screen) between three White men and three Black men. Following the conversation, 12 statements (2 made by each of the six speakers) were presented, and subjects were asked to identify who said what by matching a photo with each written statement. Category-based processing is evidenced by more within-race than between-race errors.

Inhibitory Ability. Inhibitory ability was measured by a color-naming Stroop task. Subjects were shown words (i.e., "red," "green") presented in either a red or a green hue on a black background. They were instructed to indicate the hue of each word (while ignoring the orthographic representation) by pressing a red or green key. An initial practice task was followed by 20 data-collection trials in which half the words were presented in a congruent manner (e.g., the word "red" presented in a red hue) and half were presented in an incongruent manner. Inhibitory

TABLE 2

Means and Standard Deviations for the Variables in Study 1

Variable	Nonbiased subjects		Ordinary subjects	
	Mean	*SD*	Mean	*SD*
IAT				
Difference score (ms)	88.83	112.49	320.79	158.63
d	0.22	0.20	0.73	0.28
ATB score	6.49	0.23	5.74	0.81
SAFCON task: raw ideograph ratings				
Positive priming	5.17	1.04	4.52	0.98
Negative priming	4.70	1.23	3.78	1.08
Who-said-what task: categorization errors				
Within-race errors	3.50	1.78	4.61	2.49
Between-race errors	4.22	1.59	3.41	1.51
Stroop task: color-naming latency (ms)				
Congruent trials	541.49	127.77	568.97	150.91
Incongruent trials	681.98	197.97	721.99	217.83

Note. ATB = Attitudes Toward Blacks scale; IAT = Implicit Association Test; SAFCON = susceptibility to affective conditioning.

ability was measured by subtracting reaction time on congruent trials from reaction time on incongruent trials.

Results

Manipulation Checks

Susceptibility to Affective Conditioning. For each subject, we computed the mean rating for Chinese ideographs paired with each type of picture: positive, negative, and neutral (αs = .60, .67, and .69, respectively). A one-way analysis of variance (ANOVA) with valence as the within-subjects factor was significant, $F(2, 218) = 21.18$, $p < .001$, $\eta_p^2 = .16$. Tukey's post hoc tests revealed that the ratings for each priming condition were significantly different from the ratings for the other two priming conditions (negative: $M = 3.93$, $SD = 1.15$; neutral: $M = 4.33$, $SD = .95$; positive: $M = 4.62$, $SD = 1.02$), all ps < .01, p_{rep}s > .95.

Categorization. Because there were three out-group members and only two in-group members for each speaker, we computed an adjusted score by dividing the number of within-race errors by 2 and multiplying by 3. Overall, subjects made more within-race ($M = 4.43$, $SD = 2.42$) than between-race ($M = 3.55$, $SD = 1.55$) errors, $t(109) = 2.88$, $p < .02$, $p_{rep} > .95$, $d = -0.37$.

Inhibitory Ability. As expected, reaction times were longer on incongruent trials ($M = 715.45$ ms, $SD = 214.36$ ms) than on congruent trials ($M = 564.49$ ms, $SD = 147.18$ ms), $t(109) = 12.04$, $p < .0001$, $p_{rep} > .99$, $d = 0.82$.

Prejudice Measures

The simple IAT effect was computed by subtracting the mean reaction times for compatible trials (i.e., trials on which Black names and unpleasant words, and White names and pleasant words, were paired) from the mean reaction times for incompatible trials (i.e., trials on which Black names and pleasant words, and White names and unpleasant words, were paired); the mean IAT effect was 282.49 ms ($SD = 174.49$ ms). These data were transformed into an IAT d statistic using the scoring algorithm developed by Greenwald, Nosek, and Banaji (2003); this statistic was used for all subsequent analyses involving the IAT. The mean and standard deviation of the ATB scale were 5.87 and 0.80, respectively (higher means indicate lower levels of anti-Black bias). The ATB and IAT were marginally correlated, $r(110) = -.18$, $p < .07$, $p_{rep} > .85$.

As would be expected, nonbiased and ordinary subjects differed significantly on both the implicit and the explicit measures of bias, all ps < .0001, p_{rep}s > .99 (see Table 2). In addition, nonbiased and ordinary subjects differed significantly in their ratings of ideographs in the positive-priming condition, $t(108) = 2.45$, $p < .03$, $p_{rep} > .93$, $d = 0.64$, and the negative-priming condition, $t(108) = 2.97$, $p < .01$, $p_{rep} > .95$, $d = 0.79$ (see Table 2).[2] Also, on the categorization measure, nonbiased subjects made significantly fewer within-race errors than ordinary subjects, $t(108) = 2.25$, $p = .03$, and marginally more between-race errors, $t(108) = -1.99$, $p = .059$. The two groups did not differ on inhibitory ability, $t < 1$.

To formally test the interaction between susceptibility to affective conditioning and racial bias, we computed separate scores for negative and positive affective conditioning (SAFCON$_{neg}$ and SAFCON$_{pos}$, respectively), subtracting each subject's average rating of ideographs paired with negative images from the midpoint of the rating scale (i.e., 4) and

[2]In Study 1, we obtained evidence of a positivity offset (Ito & Cacioppo, 2005): Nonbiased individuals rated ideographs in the neutral condition significantly more favorably than did ordinary subjects, $t(108) = -4.09$, $p < .001$. However, this pattern was not significant in Study 2.
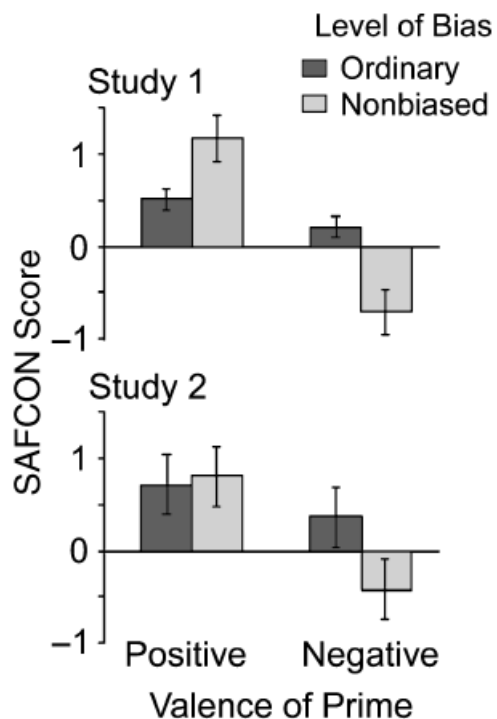
Fig. 1. Scores on the susceptibility-to-affective-conditioning (SAFCON) task as a function of valence of prime and subject group (level of bias) in Studies 1 and 2.

subtracting the midpoint of the scale from each subject's average rating of ideographs paired with positive images. Higher values indicate greater susceptibility to affective conditioning.[3] These scores were submitted to a 2 (subject group: nonbiased vs. ordinary) × 2 (valence: $SAFCON_{pos}$ vs. $SAFCON_{neg}$) mixed ANOVA with the last factor within subjects. Results revealed a significant main effect of valence, $F(1, 216) = 32.15, p < .0001, p_{rep} > .99$, $\eta_p^2 = .13$; $SAFCON_{pos}$ was greater than $SAFCON_{neg}$, indicating greater susceptibility to the acquisition of positive, compared with negative, affect overall. This main effect was qualified by a significant two-way interaction, $F(1, 216) = 16.94, p < .0001$, $p_{rep} > .99, \eta_p^2 = .07$. As shown in Figure 1, nonbiased subjects were more susceptible to positive affective conditioning than were ordinary subjects, but were significantly less susceptible to negative affective conditioning than were ordinary subjects.

To ascertain whether implicit bias, explicit bias, or the interaction of the two accounted for differences in conditioning between nonbiased and ordinary subjects, we performed two simultaneous regressions in which $SAFCON_{pos}$ and $SAFCON_{neg}$ were regressed onto centered implicit bias, centered explicit bias, and the Implicit Bias × Explicit Bias interaction.[4] Results of the analysis for $SAFCON_{neg}$, $F(3, 106) = 2.68, p = .05, R^2 =$

.07, revealed that only the interaction predicted $SAFCON_{neg}$, $\beta = -.21, p < .03, p_{rep} > .94$. For $SAFCON_{pos}$, none of the variables had a significant effect.

Discussion

The results of Study 1 demonstrate that nonbiased individuals have a general affective orientation that is distinct from that of ordinary individuals; specifically, they are less susceptible to the acquisition of negative affect and more susceptible to the acquisition of positive affect than ordinary individuals. Categorization processes also contribute to nonbias, independently of susceptibility to affective conditioning.[5] It is perhaps surprising that the results for inhibitory ability were not significant, as the Stroop task seemed to be the most relevant measure of processes related to racial associations and biases. However, inhibitory ability is a multifaceted construct (Hasher & Zacks, 1988), so it is possible that a different measure of inhibition would have been more effective. It is also important to mention that past research investigating the relation between prejudice and inhibitory ability has examined either implicit or explicit bias, whereas nonbiased individuals in the present study were defined by a combination of both implicit and explicit bias.[6]

One alternative explanation for the findings regarding affective conditioning in Study 1 is that nonbiased and ordinary subjects differed in their perceptions of the positivity or negativity of the particular conditioning stimuli used. Consequently, we used a different set of positive and negative primes in Study 2. Because research has shown that African Americans can accurately detect racial bias (Fazio et al., 1995; Livingston, 2002; Livingston, Mendoza, & Drwecki, 2007; Richeson & Shelton, 2005), we attempted to oversample nonbiased Whites by using African American nominators.

STUDY 2

Method

The subjects in this experiment were White students, 24 of whom were nominated by African Americans and 77 of whom signed up to participate through the psychology department. Subjects were classified as nonbiased or ordinary following the criteria used in Study 1. A relatively high percentage of the individuals nominated by African Americans were classified as nonbiased (see Livingston et al., 2007, for data and discussion

---

[3]We acknowledge an anonymous reviewer for pointing out that subtracting average ratings in the positive- and negative-priming conditions from average ratings in the neutral condition gives inappropriate prominence to the neutral items. Thus, we did not use this approach.

[4]All scores were linearly transformed prior to centering so that all values were greater than zero.

[5]A continuous ATB × IAT interaction term regressed onto $SAFCON_{pos}$, $SAFCON_{neg}$, $SAFCON_{pos}$ × $SAFCON_{neg}$, and a composite categorization score (within-race errors − between-race errors), $F(4, 105) = 5.43, p = .001$, yielded significant effects of all four independent variables, all $ps < .04$. In future research, we will more closely examine the complexities of how and why categorization and susceptibility to affective conditioning contribute unique variance to nonbias. However, in the remainder of this article, we focus exclusively on affective processes.

[6]Although there was no significant difference between nonbiased and ordinary subjects on inhibitory ability, $t < 1$, the simple correlation between ATB and inhibitory ability was significant, $r = .22, p < .02$. The correlation between IAT and inhibitory ability did not attain significance, $r = -.06, p = .53$.

TABLE 3

Positive and Negative Primes Used for Affective Conditioning in Studies 1 and 2

Positive primes		Negative primes	
IAPS number	IAPS name	IAPS number	IAPS name
Study 1			
2340	Family	1111	Snakes
2070	Baby	1220	Spider
1710	Puppies	1275	Roaches
Study 2			
1440	Seal	1301	Dog
1603	Butterfly	9440	Skulls
2070	Baby	2800	Sad Child

Note. IAPS = International Affective Picture System (Lang, Bradley, & Cuthbert, 2005).

TABLE 4

Means and Standard Deviations for the Variables in Study 2

Variable	Nonbiased subjects		Ordinary subjects	
	Mean	SD	Mean	SD
IAT				
Difference score (ms)	−125.75	376.30	293.21	153.14
d	0.07	0.27	0.66	0.25
ATB score	6.53	0.17	5.71	0.66
SAFCON task: raw ideograph ratings				
Positive priming	4.80	0.89	4.72	1.03
Negative priming	4.42	1.26	3.64	1.28

Note. ATB = Attitudes Toward Blacks scale; IAT = Implicit Association Test; SAFCON = susceptibility to affective conditioning.

on the accuracy of African American nominators). Three ordinary subjects were excluded from analyses because of missing data or computer malfunction. The final sample consisted of 98 individuals (12 nonbiased, 86 ordinary), who participated in exchange for extra credit or $20.

The procedure of Study 2 was very similar to that of Study 1, except that the measures of categorization and inhibition were not included, and the SAFCON task included different primes (see Table 3 for data on the positive and negative primes; as in Study 1, the neutral primes were gray squares).

Results

The one-way ANOVA on ratings of the ideographs was significant, $F(2, 190) = 30.83$, $p < .0001$, $\eta_p^2 = .25$. Tukey's post hoc tests revealed that the ratings for each priming condition differed significantly from the ratings for the other two priming conditions (negative: $M = 3.73$, $SD = 1.30$; neutral: $M = 4.05$, $SD = 0.93$; positive: $M = 4.73$, $SD = 1.01$), all ps $< .01$, p_{rep}s $> .97$ (αs $= .69, .57,$ and $.54$, respectively). The ATB and IAT were moderately correlated, $r(98) = -.28$, $p < .01$, $p_{\text{rep}} > .95$. Table 4 displays the data for ATB and IAT scores and evaluative ratings of the ideographs. As in Study 1, ordinary subjects gave lower ratings to ideographs in the negative-priming condition than nonbiased subjects did, $t(94) = 1.98$, $p = .05$, $p_{\text{rep}} > .87$, $d = 0.61$ (see Table 4). However, there were no differences in ideograph ratings in the positive-priming condition in Study 2, $t(94) = 0.28$, $d = 0.08$. As we did for Study 1, we computed SAFCON scores and submitted them to a 2 (subject group: nonbiased vs. ordinary) × 2 (valence: SAFCON$_{\text{pos}}$ vs. SAFCON$_{\text{neg}}$) mixed ANOVA. Results revealed a significant main effect of valence, $F(1, 188) = 9.78$, $p = .002$, $p_{\text{rep}} > .98$, $\eta_p^2 = .09$, indicating that subjects were more susceptible to positive than negative conditioning overall. However, this main effect was qualified by a marginally significant two-way interaction, $F(1,$

$188) = 2.98$, $p < .09$, $p_{\text{rep}} = .89$, $\eta_p^2 = .03$; although the two groups were equally susceptible to positive conditioning, ordinary subjects were more susceptible to negative conditioning than nonbiased subjects were (see Fig. 1).

As in Study 1, results of the simultaneous regression of SAFCON$_{\text{neg}}$ on implicit bias, explicit bias, and the Implicit Bias × Explicit Bias interaction, $F(3, 91) = 3.05$, $p < .04$, $R^2 = .09$, revealed that only the interaction was a significant predictor, $\beta = -.28$, $p < .02$, $p_{\text{rep}} = .96$. For SAFCON$_{\text{pos}}$, none of the independent variables had a significant effect.

GENERAL DISCUSSION

The two studies reported in this article converge in demonstrating that racial nonbias is linked to an individual's general susceptibility to the acquisition of negative versus positive affect. One widely held assumption in the literature on prejudice is that everyone in U.S. society is exposed to and aware of cultural negativity toward Blacks. It is possible that nonbiased individuals are able to avoid internalizing these negative associations to Blacks as a result of their resistance to the acquisition of negative affective associations in general. Indeed, what consistently differentiated nonbiased from ordinary subjects in both studies was that the former had a tendency to resist the acquisition of negative affective associations to neutral stimuli.

If racial bias is the product of lower-level affective processes, then it is not clear to what extent it can be successfully regulated by values or motives alone. The inability to regulate lower-level affect via higher-order reasoning may hold true for various types of attitudes. Imagine that someone with a strong aversion to lima beans learns that three servings of lima beans per day will greatly reduce the risk of certain cancers and heart disease. Although a motivated individual could alter his or her behavior to increase dietary intake of lima beans to three daily servings, it is unlikely that the same motivation would alter visceral reactions to the flavor of lima beans. Although it is possible for

individuals to "acquire" a liking for tastes that were once aversive (e.g., the taste of beer), this is almost always achieved through a process of repetitive exposure and reconditioning, rather than sheer force of will. Similarly, researchers have shown that (implicit) racial biases are not readily attenuated by egalitarian values or motives (Gaertner & Dovidio, 1986). However, these attitudes might be altered via experience-based reconditioning achieved through positive contact (Pettigrew & Tropp, 2006) or interracial dating (Olsson et al., 2005), for example.

In sum, most individuals hold racial biases not out of malice or failure to recognize the virtues of intergroup tolerance. Rather, prejudice appears to be the unfortunate consequence of lower-level affective processes that are not easily altered by higher-order reasoning. Nevertheless, it is possible that individuals without chronically low susceptibility to negative conditioning could achieve nonbias through practice, selective attention, or interpersonal experiences that gradually recondition attitudes. Indeed, Dasgupta and Greenwald (2001) demonstrated that even implicit attitudes could be improved by exposing subjects to positive, rather than negative, Black exemplars. Similarly, Olson and Fazio (2006) reduced implicit bias toward African Americans by pairing Black faces with positive stimuli. All in all, a number of studies indicate that affective, environmental, educational, and interpersonal contexts that orient individuals toward associating positive achievements or experiences with African Americans can decrease racial bias (Dasgupta & Greenwald, 2001; Ito, Chiao, Devine, Lorig, & Cacioppo, 2006; Olson & Fazio, 2006; Olsson et al., 2005; Rudman, Ashmore, & Gary, 2001). We believe that the problem of intergroup bias is most tractable when combated with reduction strategies that parallel its origins.

REFERENCES

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

Bartholow, B.D., Dickter, C.L., & Sestir, M.A. (2006). Stereotype activation and control of race bias: Cognitive control of inhibition and its impairment by alcohol. *Journal of Personality and Social Psychology*, *90*, 272–287.

Brewer, M.B., & Brown, R.J. (1998). Intergroup relations. In D.T. Gilbert, S.T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 554–594). New York: McGraw-Hill.

Brigham, J.C. (1993). College students' racial attitudes. *Journal of Applied and Social Psychology*, *23*, 1933–1967.

Cacioppo, J.T., Marshall-Goodell, B.S., Tassinary, L.G., & Petty, R.E. (1992). Rudimentary determinants of attitudes: Classical conditioning is more effective when prior knowledge about the attitude stimulus is low than high. *Journal of Experimental Social Psychology*, *28*, 207–233.

Dasgupta, N., & Greenwald, A.G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*, 800–814.

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

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.F., 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.

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.

Fiske, S.T. (1998). Stereotyping, prejudice, and discrimination. In D.T. Gilbert, S.T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 357–411). New York: McGraw-Hill.

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

Greenwald, A.G., McGhee, D.E., & Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognitions: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

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.

Hasher, L., & Zacks, R.T. (1988). Working memory, comprehension, and aging: A review and a new view. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.

Ito, T.A., & Cacioppo, J.T. (2005). Variations on a human universal: Individual differences in positivity offset and negativity bias. *Cognition and Emotion*, *19*, 1–26.

Ito, T.A., Chiao, K.W., Devine, P.G., Lorig, T.S., & Cacioppo, J.T. (2006). The influence of facial feedback on race bias. *Psychological Science*, *17*, 256–261.

Jarvis, B. (2004). DirectRT and MediaLab [Computer software]. New York: Empirisoft Corp.

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2005). *International Affective Picture System (IAPS): Technical manual and affective ratings*. Gainesville: University of Florida, Center for Research in Psychophysiology.

Livingston, R.W. (2002). Perceived negativity as a moderator of African Americans' implicit and explicit racial attitudes. *Journal of Experimental Social Psychology*, *38*, 405–413.

Livingston, R.W., Mendoza, S.A., & Drwecki, B.B. (2007). *Can Blacks tell when someone is racially biased? Accuracy in prejudice detection across multiple contexts*. Unpublished manuscript, Northwestern University, Evanston, IL.

Macrae, C.N., Milne, A.B., & Bodenhausen, G.V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, *66*, 37–47.

Monteith, M.J., & Walters, G.L. (1998). Egalitarianism, moral obligation, and prejudice-related personal standards. *Personality and Social Psychology Bulletin*, *24*, 186–199.

Nosek, B.A., Banaji, M.R., & Greenwald, A.G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration Web site. *Group Dynamics: Theory, Research, and Practice*, *6*, 101–115.

Olson, M.A., & Fazio, R.H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, *12*, 413–417.

Olson, M.A., & Fazio, R.H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, *32*, 421–433.

Olsson, A., Ebert, J.P., Banaji, M.R., & Phelps, L. (2005). The role of social groups in the persistence of learned fear. *Science*, *309*, 785–787.

Pettigrew, T.F., & Tropp, L.R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*, 751–783.

Phillips, S.T., & Ziller, R.C. (1997). Toward a theory and measure of the nature of nonprejudice. *Journal of Personality and Social Psychology*, *72*, 420–434.

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. (2005). Thin slices of racial bias. *Journal of Nonverbal Behavior*, *29*, 75–86.

Richeson, J.A., & Trawalter, S. (2005). Why do interracial interactions impair executive function? A resource depletion account. *Journal of Personality and Social Psychology*, *88*, 934–947.

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

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

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

Tajfel, H., & Turner, J.C. (1986). The social identity theory of intergroup behaviour. In S. Worchel & W.G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Chicago: Nelson-Hall.

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

von Hippel, W., Silver, L., & Lynch, M. (2000). Stereotyping against your will: The role of inhibitory ability in stereotyping and prejudice among the elderly. *Personality and Social Psychology Bulletin*, *26*, 523–532.

Zanna, M.P., Kiesler, C.A., & Pilkonis, P.A. (1970). Positive and negative attitudinal affect established by classical conditioning. *Journal of Personality and Social Psychology*, *14*, 321–328.

Zinbarg, R.E., & Mohlman, J. (1998). Individual differences in the acquisition of affectively valenced associations. *Journal of Personality and Social Psychology*, *74*, 1024–1040.