Individual Differences in the Activation and Control of Affective Race Bias as Assessed by Startle Eyeblink Response and Self-Report

David M. Amodio, Eddie Harmon-Jones, and Patricia G. Devine
University of Wisconsin—Madison

The activation and control of affective race bias were measured using startle eyeblink responses (Study 1) and self-reports (Study 2) as White American participants viewed White and Black faces. Individual differences in levels of bias were predicted using E. A. Plant and P. G. Devine's (1998) Internal and External Motivation to Respond Without Prejudice scales (IMS/EMS). Among high-IMS participants, those low in EMS exhibited less affective race bias in their blink responses than other participants. In contrast, both groups of high-IMS participants exhibited less affective race bias in self-reported responses compared with low-IMS participants. Results demonstrate individual differences in implicit affective race bias and suggest that controlled, belief-based processes are more effectively implemented in deliberative responses (e.g., self-reports).

Recent years have seen an increasing interest in the physiological correlates of social and affective processes. The incorporation of physiological measures into the study of racial prejudice has been especially appealing because these measures have the ability to circumvent response biases that often affect self-report instruments (e.g., Vanman, Paul, Ito, & Miller, 1997; Vrana & Rollock, 1998; for a review, see Guglielmi, 1999). The use of physiological measures as affective indices of racial attitudes has a relatively long history in social psychology, dating back to the mid 1950s. It was during this time in the United States that social norms prohibiting the public expression of racial prejudice began to emerge. Being wary that participants' concerns over these norms might threaten the veracity of their self-reported racial attitudes, some prejudice researchers turned to physiological indices known to be resistant to overt control efforts. This early work demonstrated that White American participants had larger autonomic responses, for example, greater skin conductance, in response to Blacks compared with Whites across a number of different experimental paradigms (e.g., Rankin & Campbell, 1955; Vidulich & Krevanick, 1966). Although this work suggested an affective race bias, the skin-conductance measure was poorly suited to differentiate the valence of White participants' affective responses to Black people.

More recent work has begun to elucidate the neural substrates of race-biased affect in order to further probe the affective qualities of intergroup bias. A large body of research has identified the amygdala, a small collection of nuclei located bilaterally in the anterior region of the temporal lobes, as playing an integral role in affective processes (Adolphs, Tranel, & Damasio, 1998; Davidson & Irwin, 1999; Whalen et al., 1998). Much research has associated amygdala activity with negative affect and, more specifically, with the detection of threat (LeDoux, 1996). Hart et al. (2000) used functional magnetic resonance imaging to measure amygdala activity in White and Black participants while they viewed pictures of White and Black faces. Greater amygdala activity was observed in participants as they viewed pictures of out-group members compared with in-group members, indicating a race bias in basic-level affective physiology.[1]

David M. Amodio, Eddie Harmon-Jones, and Patricia G. Devine, Department of Psychology, University of Wisconsin—Madison.

Portions of this research were presented at the 2000 meeting of the Society for Psychophysiological Research in San Diego, California; at the 2001 meeting of the Society for Personality and Social Psychology in San Antonio, Texas; and at the 2001 meeting of the Society for Psychophysiological Research in Montreal, Quebec, Canada. This research was supported in part by a grant from the Fetzer Institute, and by National Institute of Mental Health Emotion Training Grant T32-MH18931 for support of David M. Amodio. We thank Nikki Graf for her assistance in multiple stages of this research; Alison Covert, Donna Erhardt, Amber Peterson, and Anne Schoeneborn for their assistance in data collection and processing; Amanda Brodish and Suzanne Klonis for comments on earlier versions; and Jeremy Biesanz, Rodrigo Espana, and Alex Shackman for their helpful advice. We gratefully acknowledge the technical and mechanical support provided by Dirk Wilker, Ziggy Bialzik, Andy Mulder, and Greg Kant and the computer programming assistance provided by Craig Rypstat.

Correspondence concerning this article should be addressed to David M. Amodio, Department of Psychology, University of Wisconsin—Madison, 1202 West Johnson Street, Madison, Wisconsin 53706. E-mail: dmamodio@wisc.edu

[1] We deliberately avoid referring to race bias identified at implicit or basic affective and physiological levels as a "prejudiced" response. Although the presence of bias at these levels of processing may contribute to discriminatory behavior, it is important to distinguish such biases from explicitly held racist attitudes and beliefs (Devine, 1989; Devine, Monteith, Zuwerink, & Elliot, 1991).

Journal of Personality and Social Psychology, 2003, Vol. 84, No. 4, 738–753
DOI: 10.1037/0022-3514.84.4.738

Individual Differences in Affective Race Bias

Using functional magnetic resonance imaging, Phelps et al. (2000) examined the potential moderating role of prejudice level on amygdala activity in White participants as they viewed pictures of Blacks and Whites. Phelps et al. did not observe a reliable relationship between participants' amygdala activity and scores on McConahay's (1986) Modern Racism Scale (MRS), a self-report, explicit measure of racially prejudiced attitudes. They concluded that the amygdala activity indexed an implicit (unconscious) level of race bias not evident in self-reported racial attitudes. In support of this interpretation, they found that amygdala activity level, but not MRS score, was correlated with two other indices of implicit race bias: (a) the affect-modulated startle eyeblink response during presentations of Black faces and (b) the Implicit Association Test (IAT, A. Greenwald, McGhee, & Schwartz, 1998), a reaction-time measure of the evaluative categorization of Black and White faces.

Phelps et al.'s (2000) results indicate that there is variability in individuals' prejudice-related affective responses that is not being accounted for by self-reported attitude measures of prejudice. Like much recent work assessing implicit forms of race bias, these results suggest that not all people are equally prone to implicit race bias (e.g., Fazio, Jackson, Dunton, & Williams, 1995; Lepore & Brown, 1997; Moskowitz, Saloman, & Taylor, 2000; Wittenbrink, Judd, & Park, 1997). Devine, Plant, Amodio, Harmon-Jones, and Vance (2002) found that some individuals who report low prejudice on attitude measures exhibit relatively high levels of implicit race bias, and they suggested that this inconsistency contributes to the low correspondence between attitude measures and implicit assessments. A major goal of the present research is to examine the differences in automatic levels of affective race bias that exist among low-prejudice individuals. In the next section, we briefly review individual difference factors that have been shown to moderate levels of implicit race bias in recent work reported by Devine et al. (2002).

Internal and External Motivation to Respond Without Prejudice

Although the earlier research of Devine and her colleagues focused rather exclusively on the nature of people's intrapersonal struggles to overcome prejudice (e.g., Devine, 1989; Devine & Monteith, 1993; Devine, Monteith, Zuwerink, & Elliot, 1991), their more recent theorizing has acknowledged the additional importance of interpersonal concerns and normative influences proscribing prejudice in contemporary American society (Devine et al., 2002; Plant & Devine, 1998, 2001). Plant and Devine (1998) introduced the Internal and External Motivation to Respond Without Prejudice scales (IMS and EMS, respectively) to measure the normative as well as the personal motivations that lead people to control expressions of prejudice. The IMS measures the extent to which nonprejudiced responses are motivated by individuals' personal beliefs and values and includes items such as "I attempt to act in nonprejudiced ways toward Black people because it is personally important to me" and "Being nonprejudiced toward Black people is important to my self-concept." The EMS measures the extent to which nonprejudiced responses are motivated by external forces and includes items such as "If I acted prejudiced toward Black people, I would be concerned that others would be angry with me" and "I attempt to appear nonprejudiced toward Black people in order to avoid disapproval from others."[2]

Plant and Devine (1998) demonstrated the IMS and EMS to be reliable and provided evidence regarding the scales' convergent, discriminant, and predictive validity. For example, they demonstrated that scores on the IMS were highly correlated with self-report measures of prejudice, including the Attitude Towards Blacks scale (ATB, Brigham, 1993) and the MRS, such that higher levels of internal motivation were associated with lower prejudice scores. The EMS, in contrast, was only modestly correlated with prejudice measures, such that high levels of external motivation were weakly associated with high prejudice scores. In addition, only a small correlation was found between the EMS and measures of general social evaluation such as the Interaction Anxiousness Scale (Leary, 1983), the Marlowe–Crowne Social Desirability Scale (Crowne & Marlowe, 1960), and Snyder and Gangestad's (1986) Self-Monitoring Scale. The IMS and EMS have been shown to be largely independent (average $r = -.14$ across multiple samples). Thus, individuals can be motivated to respond without prejudice primarily for internal reasons, primarily for external reasons, for both internal and external reasons, or they may not be motivated for either type of reason.

An important implication of the IMS–EMS conceptualization is that it distinguishes between two classes of people who are personally motivated to respond without prejudice: those with and those without external motivations to respond without prejudice. In support of this distinction, the program of research developed by Devine, Plant, and their colleagues (Devine et al., 2002; Plant & Devine, 2000, 2002) has demonstrated that high-IMS, low-EMS participants and high-IMS, high-EMS participants differ substantially in their self-reported and actual abilities to respond without prejudice and in their levels of implicit race bias. Although one might generally expect that being strongly motivated for both internal and external reasons would be associated with more effective control of prejudice than being motivated for only internal reasons, research and theory suggest that this is not the case. Rather, individuals who are primarily internally motivated are more effective in responding without prejudice than those motivated for both reasons. For example, Plant and Devine (2000) found that high-IMS, low-EMS participants were more confident in their self-reported ability to respond without prejudice than high-IMS, high-EMS participants. They reported not being nervous when approaching intergroup interactions, whereas high-IMS, high-EMS participants reported higher levels of nervousness when approaching intergroup interactions. When provided with an opportunity to engage in a task (computer program) that would help to alleviate covert bias (i.e., bias not observable to others), only high-IMS, high-EMS participants spent large amounts of time on the task (Plant & Devine, 2002). Consistent with their self-perceived efficacy, participants high in IMS but low in EMS showed little interest in the bias-reduction computer program.

---

[2] Internal and external motivation to respond without prejudice scales have now been developed for motivation to respond without sexism (Klonis & Devine, 2000), homophobia (K. Lemm, personal communication, August 11, 1999), and prejudice toward fat people (Buswell & Devine, 2000). In all cases, the internal and external scales were independent.

Finally, Devine et al. (2002) compared levels of implicit race bias with self-reported attitudes toward Blacks among individuals with varying levels of IMS and EMS. They found that whereas all high-IMS participants reported similarly positive attitudes toward Black people on the ATB, high-IMS, low-EMS participants exhibited significantly lower levels of implicit race bias than high-IMS, high-EMS participants on two alternative implicit measures (sequential evaluative priming and the IAT).[3]

The pattern of differences between high-IMS, low-EMS participants and high-IMS, high-EMS participants is consistent with motivation and goals research, such that the pursuit of a particular goal is effective to the extent that it has been internalized into one's self-concept. Self-determination theory, for example, posits that the motivations for people's responses may be placed along a continuum from the primarily external to the primarily internal (Deci & Ryan, 2000; Ryan & Connell, 1989). For example, at the lowest end of the self-determination continuum, responses are characterized as being primarily externally motivated and are expected to occur only in the presence of others. Moderately self-determined responses are characterized as representing a combination of external and internal motivations. According to Deci and Ryan (2000), these moderately self-determined responses are better maintained than less self-determined responses, yet they are unstable and not effectively regulated across situations. Finally, responses that are highly self-determined are characterized as being primarily internally motivated. These highly self-determined responses arise from well-elaborated, internalized processes and are effectively regulated across situations. In support of this overarching conceptualization, research shows that behavior motivated out of more self-determined (i.e., high internal and low external) reasons leads to more effective strategies for goal attainment and greater long-term efficacy of self-regulatory efforts (e.g., Grolnick & Ryan, 1987; Ryan, Rigby, & King, 1993; Williams, Rodin, Ryan, Grolnick, & Deci, 1998), as well as greater attitude–behavior consistency (Koestner, Bernieri, & Zukerman, 1992). This work is consistent with the findings of Devine et al. (2002), in which participants who were primarily internally motivated (high IMS, low EMS) responded without prejudice more effectively on difficult-to-control implicit measures compared with participants motivated for both internal and external reasons (high IMS, high EMS).

The cumulative program of recent research by Devine, Plant, and their colleagues has suggested that a consideration of both internal and external motivations to respond without prejudice is crucial when examining individual differences in the activation and control of race bias. Although all high-IMS individuals share egalitarian beliefs, those high in EMS are more prone to transgressing their nonprejudiced standards in their expected and actual behaviors. These findings demonstrate that the internalization of nonprejudiced standards has important implications for people's ability to effectively respond without prejudice across a range of responses. However, the effect of such internalization, as assessed by the IMS and EMS, on the activation and control over the affective component of race bias has not been examined previously. Indeed, Hamilton (1981) noted that "if there is any domain of human interaction that history tells us is laden with strong, even passionate, feelings, it is in the area of intergroup relations" (p. 347). Examining the implications of IMS and EMS on the activation and control of affective race bias represents an important but as yet unstudied area of prejudice research.

Study 1

Our goal in Study 1 was to examine individual differences in affective race bias during the very early stages of processing when automatic effects occur and also at later times when control may be possible. To this end, we built on the recent work of Devine et al. (2002) to examine the moderating effects of IMS and EMS on a physiological index of affective race bias, namely, startle eyeblink modulation in response to Black compared with White faces. A major appeal of the startle eyeblink measure is that eyeblink magnitude can be precisely assessed at any time while participants are viewing stimuli, and therefore the measure possesses extremely high temporal resolution. Consequently, when used to assess short latency responses, the startle eyeblink measure can reveal affective responses that are clearly attributable to automatic processes. By comparison, the extent to which conventional reaction-time-based measures can reveal automatic race bias is somewhat ambiguous because the time taken to deliver responses can vary widely, and therefore these measures do not necessarily preclude controlled processing (Payne, 2001). Thus, an additional goal of Study 1 was to examine automatic affective race bias using a measure that more effectively precludes controlled processes than reaction-time measures.

Startle Eyeblink Modulation

Startle eyeblink modulation has been used to measure affective processes in a large corpus of research (for a review, see Bradley, Cuthbert, & Lang, 1999). The eyeblink is a component of the whole-body startle reflex, and its magnitude can be measured by electrodes placed over the muscle below the eye. Numerous studies have demonstrated that negative affective states amplify startle reactions, whereas positive affective states inhibit them, shortly after the onset of affective stimuli (e.g., Filion, Dawson, & Schell, 1998; Lang, Bradley, & Cuthbert, 1990; Sutton, Davidson, Donzella, Irwin, & Dottl, 1997). Affective dispositions toward objects can be assessed by presenting a startle probe (short blast of white noise that causes an eyeblink) to participants while they are viewing the object and determining the magnitude of the resulting blink. Thus, the modulation of the startle eyeblink response represents the summation of affective responses to the probe and the affective foreground stimulus (e.g., an emotional picture).

Previous research on the time course of nonphysiological race-biased reactions has demonstrated that responses measured at early stages in the processing of social stimuli reflect automatic levels of

---

[3] As discussed by Devine et al. (2002), the finding that some high-IMS individuals exhibit less race bias on implicit measures than others suggests a revision to Devine's earlier work (e.g., Devine, 1989), such that although on average all high-IMS individuals may possess a certain level of implicit bias, this level of bias is moderated by EMS. Indeed, research examining the degree to which high-IMS, low-EMS participants possess significant levels of implicit race bias has to date yielded mixed results, with evidence of bias provided by some studies but not by others (see Devine et al., 2002). Thus, the findings of Devine et al. (2002) represent an extension of Devine (1989) rather than a disparity with the earlier work.

bias, whereas responses measured later reflect the potential involvement of control (Devine, 1989; Kawakami, Dion, & Dovidio, 1998). By measuring eyeblink responses to startle probes occurring at short and long latencies following the onset of Black compared with White faces, we were able to examine affective processes associated with both the activation and potential control of race bias.

Importantly, patterns of eyeblink modulation have different meanings at short and long latencies. Eyeblink responses at short latencies (less than 500 ms following stimulus onset) vary as a function of attention. An aroused attentional state is evoked by the perception of affectively relevant stimuli, either positive or negative in valence (Bradley et al., 1999; Öhman, Flykt, & Esteves, 2001). The consequence of this state is the inhibition of the startle response (Bradley, Cuthbert, & Lang, 1993; Vanman, Boehmelt, Dawson, & Schell, 1996). A race-biased response to a short latency probe would be evidenced by relatively smaller blinks (blink *inhibition*). Although blink inhibition to short latency probes may theoretically be associated with either positive or negative affect, converging theory and evidence based on other data suggest that the initial affective response of a White American participant toward a Black individual would be negative (e.g., Devine, 1989; Devine et al., 2002; Fazio et al., 1995, Hart et al., 2000; Phelps et al., 2000; Wittenbrink et al., 1997). Startle eyeblinks at long latencies (500 ms or greater) are modulated as a function of affective valence, such that blink inhibition is associated with a positive affective response, and blink *amplification* is associated with a negative affective response (e.g., Lang et al., 1990). That is, a negative race-biased affective response to a long latency probe would be evidenced by larger amplitude blinks. Although short and long latency blink responses vary as a function of attention and affective valence, respectively, both arise from underlying affect-related processes, and thus both reveal basic affective responses.

Hypotheses

Our hypotheses for individual differences in White American participants' basic affective responses to Black and White faces follow directly from the theorizing of Devine et al. (2002). Consider first the hypothesis for automatic affective responses, as measured by short latency startle eyeblinks to Black faces. Recall that Devine et al. (2002) found that the primarily internally motivated participants (high IMS, low EMS) exhibited lower levels of implicit race bias than all other participants. On the basis of this work, we predicted that high-IMS, high-EMS participants and low-IMS participants would exhibit higher levels of automatic affective race bias, as indicated by blink inhibition to Black faces at 400 ms, compared with high-IMS, low-EMS participants. This prediction assumes that the low levels of implicit bias observed for high-IMS, low-EMS participants by Devine et al. (2002) is a result of lower race-bias activation rather than highly adept control. As described above, the reaction-time measures used by Devine et al. (2002) allowed for the possibility that race bias was initially activated and then quickly controlled. By comparison, the capacity for control among high-IMS, low-EMS participants would be severely limited during short latency blink responses to Black faces, and therefore the present study's use of startle eyeblink methodology provides a stronger test of individual differences in race-bias activation levels than is typically available from reaction-time measures.

Our second set of hypotheses concerns startle eyeblink responses to Black faces at long latencies, when controlled processes are theoretically available. Devine (1989) theorized that the initial activation of race bias among low-prejudice individuals is subsequently controlled and replaced with belief-consistent nonprejudiced responses. However, the extent to which controlled processes operate in basic affective race bias has not previously been examined. Two hypotheses are plausible. The first hypothesis is that controlled processes do not operate in basic affective responses, and therefore the pattern of affective response would not change from the short to the long latency eyeblink measures. Following from our short latency blink predictions, we would expect high-IMS, high-EMS participants and low-IMS participants to exhibit greater negative affect, indicated by larger blinks at long latencies, compared with high-IMS, low-EMS participants. The alternative hypothesis assumes that controlled processes modulate basic level affect, such that long latency blinks would be amplified only for low-IMS participants in response to Black faces and that blink amplification would not be evident in the responses of either high-IMS group (i.e., high IMS, low EMS or high IMS, high EMS).

The design of Study 1 allowed us to examine some plausible alternative explanations for our predicted results. First, to address the possibility that individual differences in responses to Black faces could be due to general out-group negativity, we included faces of another minority group: Asians. By demonstrating that the sources of motivation to respond without prejudice toward Black faces do not also predict differential response to Asian faces, we can be confident in the specificity of our finding. A separate, second concern is that responses toward Black faces could be due to a general unfamiliarity with minority group members among our White participants. This concern would be ruled out by demonstrating motivation group effects for Black but not Asian faces, provided that Black and Asian faces are similarly unfamiliar to our White participants. A third concern addressed in our design is that individuals with varying levels of IMS and EMS may differ in their responses to general affective stimuli unrelated to race. To address this possibility, we presented general affective pictures from the International Affective Picture System (IAPS; Center for the Study of Emotion and Attention, 1995) in a second block of trials. Finally, we obtained participants' responses to the ATB to compare the predictive ability of the IMS–EMS with a traditional attitude measure of prejudice.

Method

Participants. Sixty-seven right-handed White female students enrolled in Introductory Psychology participated voluntarily for extra course credit. As in past research on emotion, a female-only sample was used to reduce possible sex-related variability in physiological responses, and right-handed participants were selected to avoid physiological differences due to brain laterality (e.g., Davidson, Ekman, Saron, Senulis, & Friesen, 1990; Ekman, Davidson, & Friesen, 1990). The IMS ($\alpha = .81$), EMS ($\alpha = .82$), and ATB ($\alpha = .88$) were completed during a mass testing session held early in the semester. Responses to the IMS and EMS were given on a 9-point scale ranging from 1 (*strongly disagree*) to 9 (*strongly agree*). A complete list of IMS and EMS items, along with detailed descriptions of their psychometric properties, may be found in Plant and Devine (1998).

Responses to the ATB were given on a 7-point scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). After appropriate reverse scoring of some items, participants' responses were averaged to produce the final ATB score, with higher scores indicating lower levels of prejudice. Experimenters were blind to participants' scores on these individual difference measures.

Correlation analyses of IMS, EMS, and ATB scores were performed on responses from the two mass testing samples from which participants were recruited. Because significance levels are inflated by the large sample size, effects should be interpreted according to Cohen and Cohen's (1983) recommendations, such that correlations of .10 are small, .30 are moderate, and .50 are large. Analyses revealed the correlation between IMS and EMS scores to be small, $r(2552) = -.07$, $p = .001$. The correlation between ATB and IMS scores was large, $r(2552) = .65$, $p < .001$, and the correlation between ATB and EMS was moderate, $r(2552) = -.26$, $p < .001$. This pattern of intercorrelations is consistent with past research (Devine et al., 2002; Plant & Devine, 1998, 2001; Plant, Devine, & Brazy, 2003).

As in previous work, we selected participants from the extreme ends of the IMS and EMS distributions (Devine et al., 2002; Plant & Devine, 2001). The participants comprised a high-IMS, low-EMS group ($M_{\text{IMS}} = 8.84$, $M_{\text{EMS}} = 2.42$), a high-IMS, high-EMS group ($M_{\text{IMS}} = 8.79$, $M_{\text{EMS}} = 7.14$), and a low-IMS group ($M_{\text{IMS}} = 6.49$, $M_{\text{EMS}} = 4.50$). The low-IMS participants were examined as a single group for theoretical reasons. Low-IMS individuals, regardless of their level of EMS, respond with relatively high levels of prejudice on traditional self-report scales and have been shown in previous research to exhibit high levels of bias on implicit measures (Devine et al., 2002). Given that past research has shown a correlation between implicit and physiological measures of race bias (Phelps et al., 2000), these individuals would not be expected to vary in their startle eyeblink responses. In addition, situational factors that have been found to moderate low-IMS participants' responses on some measures, such as whether responses are made privately or publicly (Amodio & Devine, 2000; Plant & Devine, 1998, 2001), were carefully controlled and not manipulated in the present study. Therefore, low-IMS, high-EMS participants and low-IMS, low-EMS participants were theorized to respond similarly across measures in the present study, and preliminary analyses of blink responses did not reveal statistical differences on the primary dependent measures.

Design. Participants from each motivation group viewed pictures of Black and White faces. For each picture type, four trials contained a startle probe occurring 400 ms after picture onset, four contained a startle probe occurring 4,000 ms after picture onset, and four contained no startle probe (i.e., clear trial). Hence, the basic experimental design was a 3 (motivation group: high IMS, low EMS vs. high IMS, high EMS vs. low IMS) × 2 (probe latency: 400 ms vs. 4,000 ms) × 2 (race of face: Black vs. White) mixed factorial, with repeated measures on probe latency and race of face. Four startle probes were presented during intertrial intervals (ITIs) to enhance participants' perception that startle probes were occurring randomly. Trials were presented in a quasi-random order, with no single trial type repeated more than once consecutively. To address possible alternative explanations of our predicted results, Asian faces were included in the face trials and a separate block of trials was included to assess responses to positive, neutral, and negative IAPS pictures.

Procedure. Participants were run individually. After providing informed consent, participants were prepared for physiological recording.[4] The experimenter explained that we were interested in measuring brainwave responses to different pictures (Harmon-Jones & Allen, 2001), and that the first set of pictures contained faces and the second set contained a variety of objects, ranging in affective quality from pleasant to disturbing. Participants were instructed to pay attention to each picture throughout its duration without responding verbally. To familiarize participants with the general procedure and the sound of the startle probe, four neutral IAPS picture trials preceded the experimental trials, with startle probes occurring during three pictures and one probe occurring during an ITI. Participants then viewed the face pictures in two sets, separated by a 1-min break, and then viewed the IAPS pictures in two sets, also separated by a 1-min break. Pictures were presented for 6 s, and ITIs ranged from 14 s to 22 s in length. A fixation point appeared in the center of the computer screen 3 s prior to the onset of each picture. Presentation of face trials and IAPS trials lasted approximately 40 min. At the end of the experimental session, participants were debriefed, thanked, and dismissed.

Materials. Stimuli used in the first set of trials were pictures of male faces varying in race (12 White, 12 Black, and 12 Asian), each with a neutral expression, digitized at 640 × 500 pixels. Stimuli in the second set of trials were pictures from the IAPS (12 positive, 12 neutral, 12 negative), selected based on normative ratings of valence and arousal.[5] Pictures were presented on a 19-inch (48.26 cm) computer monitor situated approximately 3 ft (0.91 m) in front of the participant. An acoustic startle probe, consisting of a digitized 50-ms burst of white noise at 96 dB with near instantaneous rise time, was generated by a sound card (16 Sound Blaster, Creative Labs, Milpitas, CA), amplified by an audio receiver (Sony STR-DE310, New York, NY), and presented binaurally through stereo headphones (Sony MDR-CD60, New York, NY). Probe volume was initially calibrated using a precision sound-level meter (Brüel and Kjær 2203, Nærum, Denmark) with the appropriate headphone coupler and calibrated prior to each experimental session using a different sound level-meter (Radio Shack 33-2055, Fort Worth, TX). Stimulus presentation was performed using DMDX software, developed at Monash University and at the University of Arizona (Forster & Forster, in press).

Pretests of materials. In two separate pretests, students from an introductory psychology course rated the Black, White, and Asian faces used in Studies 1 and 2 on attractiveness and familiarity. In both pretests, participants viewed each picture for 6 s and then made their rating.

In the pretest examining attractiveness, 63 participants were asked to rate how attractive each face was "in an objective sense, and not how attracted you may feel toward it," on a scale ranging from 1 (*very unattractive*) to 10 (*very attractive*). Ratings were averaged within race of face and submitted to a 3 (motivation group: high IMS, low EMS vs. high IMS, high EMS vs. low IMS) × 3 (race of face: White vs. Black vs. Asian) mixed-factor analysis of variance (ANOVA), with repeated measures on race of face. This analysis produced a main effect for race of face, $F(2, 120) = 49.85$, $p < .001$, $r = .54$. Post hoc tests using Fisher's least significant difference (LSD) test revealed that this large effect was driven by lower attractiveness ratings for Asian faces ($M = 3.66$, $SD = 1.01$) compared with White faces ($M = 4.66$, $SD = 0.83$), $t(61) = 9.14$, $p < .001$, $r = .76$, and Black faces ($M = 4.53$, $SD = 0.97$), $t(61) = 8.00$, $p < .001$, $r = .72$. Most importantly, attractiveness ratings of White and Black faces did not differ significantly,

---

[4] Other physiological measures, including electroencephalogram (EEG), were collected in the experimental sessions but are not relevant to the present theoretical analysis and therefore are not reported. The only measure taken prior to those described in Study 1 was of resting EEG for purposes of comparing prefrontal resting asymmetry to self-report measures unrelated to prejudice, as in past research (e.g., Harmon-Jones & Allen, 1997, 1998). Indices of prefrontal asymmetry (right–left side alpha band activity, collected with eyes open and closed) did not differ significantly between motivation groups ($F\text{s} < 1$).

[5] The following IAPS pictures were selected on the basis of ratings of valence and arousal published by Lang et al. (1997): Positive slides 1440, 1463, 1650, 5621, 1811, 2040, 7270, 7230, 7330, 8030, 8080, 8501; neutral slides 5500, 5740, 7000, 7002, 7010, 7090, 7130, 7170, 7500, 7560, 7950; negative slides 1120, 1220, 1270, 1300, 1930, 3060, 3110, 3350, 6230, 9300, 9410, 9570. Mean ratings of valence and arousal for the pictures used in this study were 7.60 and 5.77 for positive pictures, 5.05 and 2.97 for neutral pictures, and 2.63 and 6.38 for negative pictures, respectively, scored on a scale ranging from 1 (*lowest*) to 9 (*highest*).

$t(61) = 1.20$, $p = .23$, $r = .15$ (95% confidence interval [CI] $= -.05, .20$). These results suggest that any differences in affective responses to the White and Black faces used in Study 1 would not be due to differences in attractiveness. In addition, significant effects were not obtained for motivation group, $F(2, 60) = 1.07$, $p = .35$, $r = .14$, or the Motivation Group × Race of Face interaction, $F(4, 120) = 0.89$, $p = .47$, $r = .09$.

In the pretest examining familiarity, 18 participants rated each face according to how familiar it was on a scale ranging from 1 (*very unfamiliar*) to 10 (*very familiar*). The experimenter told participants that although they may never have seen these faces before, some may seem very familiar whereas others may not seem familiar at all. Ratings were averaged within race type. A one-way within-subjects ANOVA conducted on familiarity ratings of Black versus White versus Asian faces was significant, $F(2, 34) = 8.89$, $p < .005$, $r = .46$. Fisher's LSD analyses revealed that White faces were rated as significantly more familiar ($M = 4.58$, $SD = 1.00$) than Black faces ($M = 4.06$, $SD = 1.31$), $t(17) = 2.45$, $p < .05$, $r = .51$, and Asian faces ($M = 3.64$, $SD = 1.13$), $t(17) = 4.11$, $p < .01$, $r = .71$. However, familiarity ratings did not differ significantly for Black and Asian faces, $t(17) = 1.83$, $p = .09$, $r = .41$ (95% CI $= -.03, .42$), although Asian faces were judged to be marginally less familiar.

Startle eyeblink recording and analysis. To record eyeblinks, 4 mm Ag/AgCl electrodes (In Vivo Metric, Healdsburg, CA) were placed over the left inferior orbicularis oculi below the inner and outer canthi, as suggested by van Boxtel, Boelhowser, and Bos (1998). ECI Electro-Gel (Eaton, OH) was used as the conductive medium. A mild abrasive was applied to the skin before electrode application to lower impedance levels to 10 kOhms or below (Harmon-Jones & Allen, 2001). The raw electromyographic (EMG) signal was amplified (20,000 times), and frequencies below 30 Hz and above 500 Hz were filtered online (Contact Precision Instruments Bio II, Cambridge, MA). Signals were digitally sampled at 2 kHz for 1 s prior to and throughout the duration of the trial. A separate channel was used to record the startle stimulus in order to verify its onset latency for each trial. Startle eyeblink amplitude was determined by calculating the root mean square of EMG signal between 30 ms and 90 ms following probe onset (Grillon & Davis, 1995).

Removal of EMG artifact (e.g., extraneous movement) excluded 2.52% of the total trials for the face block and 8.47% for the IAPS block. Artifact-free blink responses were standardized to *z* scores within participants, and then transformed to *t* scores, producing a distribution with a mean of 50 and a standard deviation of 10, as in Globisch, Hamm, Esteves, and Öhman (1999). Blink amplitudes greater than three standard deviations from the mean within the distributions of individual participants were considered to be outliers and were omitted from analyses. Outliers accounted for an additional 1.17% of the trials for the face block and 2.13% for the IAPS block. Acceptable blink scores were averaged within trial types to yield blink amplitude scores for each picture type at each probe latency.

Results

Data from 8 of the original 67 participants were excluded from analysis. These omissions resulted from equipment failure (3 participants), experimenter error in data collection (3 participants), participant's termination of the session prior to the experimental trials (1 participant), or because average blink scores for a condition were greater than three standard deviations from the mean (2 participants). This resulted in the exclusion of data from 2 high-IMS, low-EMS participants, 2 high-IMS, high-EMS participants, and 3 low-IMS participants. An additional participant was excluded because of incomplete IMS–EMS data.

Data analysis strategy. Participants were selected from the extremes of the IMS and EMS distributions. Extreme group designs are commonly used to strengthen the internal validity of experiments and are recommended when participant attributes (e.g., high prejudice or low IMS) that are crucial to hypothesis testing are poorly represented in participant populations (Alloy, Abramson, Raniere, & Dyller, 1999). This strategy is particularly common in psychopathology research (e.g., Dikman & Allen, 2000; Showers, Abramson, & Hogan, 1998). However, because the selection of extreme groups renders a non-normal sample distribution, treating their scores as continuous measures is not appropriate. Hence, our extreme group data were analyzed using ANOVA. On the basis of our hypotheses for individual differences, we examined blink responses using planned comparisons, whereby high-IMS, low-EMS participants were contrasted with the combination of high-IMS, high-EMS participants and low-IMS participants. Thus, these were tested using 2 (motivation group comparison: high IMS, low EMS vs. high IMS, high EMS and low IMS) × 2 (race of face: Black vs. White) mixed-factorial ANOVAs, with repeated measures on race of face. Because all a priori comparisons were derived from theory and were directional, they were evaluated using a one-tailed criterion of significance (Hayes, 1988; Rosenthal, Rosnow, & Rubin, 2000).

Responses at the 400-ms and 4,000-ms probe latencies were analyzed separately because they index conceptually different processes (i.e., attentional modulation vs. affective modulation). Although both are associated with activation of the startle pathway and affective processes, attentional and affective startle modulation responses are quantified on different theoretical scales. Indices of internal reliability for startle responses were characteristically low (Cronbach's alphas ranged between .23 and .32), although the alphas were similar across conditions. The low reliability of the startle eyeblink index is due to the measure's high level of sensitivity to changes in attention and affect (Hawk & Cook, 2000). However, low internal reliability has not been considered to be a problem for startle eyeblink measures of affect, because affect-modulated startle responses have been demonstrated to possess acceptable test–retest reliability (Larson, Ruffalo, Nietert, & Davidson, 2000) and to reliably assess changes in affective responses across many experiments (e.g., Bradley et al., 1999). Moreover, because lower internal reliability contributes to random measurement error (e.g., noise), tests of startle eyeblink effects are on average more conservative and tend to yield inflated confidence intervals.

Finally, effect sizes are presented as recommended by Rosenthal et al. (2000). Effect sizes are reported as the point-biserial correlation (r_{pb}), which indicates the magnitude of a manipulation's effect on participants' responses. The effect size *r* is interpreted as a Pearson product–moment correlation, ranging from -1 to $+1$, with positive values indicating an effect in the predicted direction. Effect size *r* values of .10, .30, and .50 correspond to Cohen's *d* of .20, .63, and 1.15, respectively. In addition, 95% CIs for the effect sizes are reported for null effects that were consistent with hypotheses.

Short latency startle responses. A 2 (motivation group comparison: high IMS, low EMS vs. high IMS, high EMS and low IMS) × 2 (race of face: Black vs. White) mixed-factorial ANOVA, with repeated measures on race of face, was first conducted on blink responses to the 400-ms probe. This analysis produced only a main effect for the motivation group comparison, $F(1, 56) = 5.15$, $p < .03$, $r = .29$. As can be seen in Figure 1, there was a general tendency among high-IMS, high-EMS participants

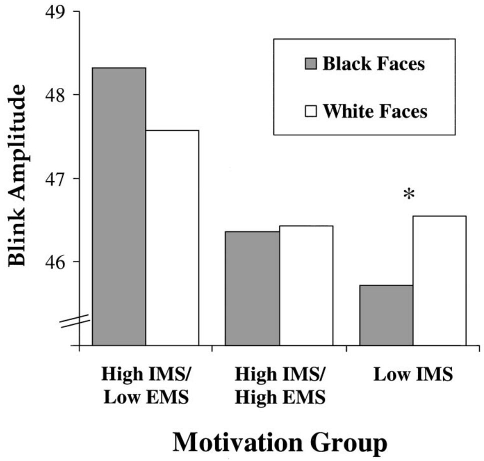

Figure 1. Mean startle eyeblink amplitude in response to Black and White faces at the 400-ms startle probe latency, as a function of motivation group. Blink amplitude is represented in *t* scores. Lower blink amplitude indicates greater affective activation. An asterisk denotes a significant difference ($p < .05$) in Black versus White responses within motivation group. Between-subjects differences are reported in the text. IMS = Internal Motivation Scale; EMS = External Motivation Scale.

and low-IMS participants to be more attentive to faces of both races, perhaps indicating a heightened vigilance to race-related stimuli in general. The relatively greater vigilance of these participants may have masked a Motivation Group Comparison × Race of Face interaction.

Because our theorizing made the specific a priori prediction that high-IMS, low-EMS participants would evidence less blink modulation than other participants to Black faces but not to White faces, we examined the effects of the motivation group comparison on Black faces and White faces separately. In support of our hypothesis, blink amplitudes to Black faces for the high-IMS, high-EMS participants and low-IMS participants ($M = 46.04$, $SD = 3.11$) were smaller than those of the high-IMS, low-EMS participants ($M = 48.32$, $SD = 5.17$), $t(56) = 2.10$, $p < .02$, $r = .27$. This comparison did not produce a significant effect for White faces, $t(56) = 1.16$, $p = .25$, $r = .15$ (95% CI $= -.39, .84$). Because the patterns of participants' startle eyeblinks to Black and White faces were similar (see Figure 1), we repeated the motivation group comparison for blink responses to Black faces after responses to White faces had been covaried using a regression procedure. The comparison remained significant, $t(56) = 1.97$, $p < .03$, $r = .25$, suggesting a motivation group effect for Black faces over and above any differences in responses to White faces. Follow-up analyses of blink responses to Black faces tested each pairwise comparison between the three motivation groups. These showed that, compared with high-IMS, low-EMS participants, blink amplitudes were significantly smaller among high-IMS, high-EMS participants ($M = 46.36$, $SD = 3.04$), $t(56) = 1.80$, $p < .02$, $r = .23$, and low-IMS participants ($M = 45.72$, $SD = 3.24$), $t(56) = 2.39$, $p < .01$, $r = .30$. Blink responses of high-IMS, high-EMS participants and low-IMS participants were not significantly different, $t(56) = 0.59$, $p = .29$, $r = .08$ (95% CI $= -.33, .79$), although the large CI prevents us from concluding strongly that a difference does not truly exist between these groups.

Across analyses, these results support the prediction that basic-level affective processes are activated to a lesser degree for high-IMS, low-EMS participants compared with other participant groups. Although theory and past research suggest that the affective responses toward Black faces would have been negative, attention-modulated startle responses do not by themselves specify the direction of affective valence. Overall, our interpretation of these findings must be qualified by the fact that the Motivation Group × Race interaction was not significant. Furthermore, although the motivation group comparison effect for responses to Black faces is suggestive of affective race bias at 400 ms, the large CIs reported for nonsignificant effects limit our ability to infer that motivation groups did not differ in their responses to White and Asian faces. It should be noted, however, that these nonsignificant effects were obtained in the context of significant differences for responses to Black faces, and because the characteristically low internal reliability of the startle eyeblink measure contributes to an inflated confidence interval, some caution is warranted when interpreting the meanings of these CIs.

Long latency startle responses. A 2 (motivation group comparison: high IMS, low EMS vs. high IMS, high EMS and low IMS) × 2 (race of face: Black vs. White) mixed-factorial ANOVA, with repeated measures on race of face, was conducted on 4,000-ms blink responses to Black and White faces. This analysis produced a significant effect for race of face, $F(1, 56) = 10.95$, $p < .001$, $r = .40$, and a marginal effect for the motivation group comparison, $F(1, 56) = 3.55$, $p < .07$, $r = .24$. These main effects were qualified by a significant interaction, $F(1, 56) = 7.98$, $p < .01$, $r = .35$. The average blink responses to Black and White faces for each motivation group are illustrated in Figure 2.

A series of planned comparisons tested our competing predictions for individual differences in affective race bias when controlled processes were theoretically available. The first hypothesis was that eyeblinks at long latencies reflect uncontrolled responses,

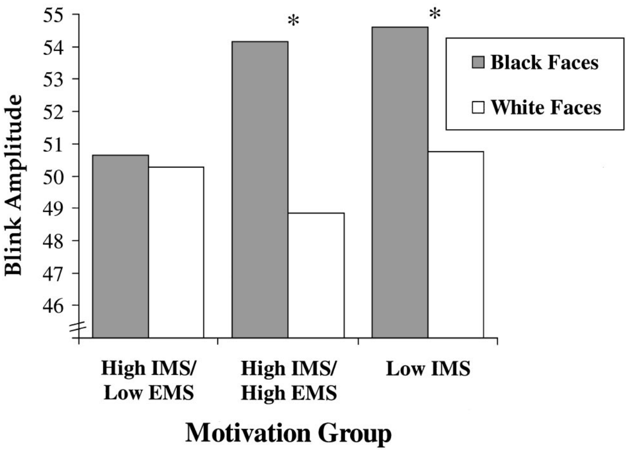

Figure 2. Mean startle eyeblink amplitude in response to Black and White faces at the 4,000-ms startle probe latency, as a function of motivation group. Blink amplitude is represented in *t* scores. An asterisk denotes a significant difference ($p < .05$) in Black versus White responses within motivation group. Between-subjects differences are reported in the text. IMS = Internal Motivation Scale; EMS = External Motivation Scale.

similar to responses at short latencies, such that blink amplitudes to Blacks would be lower for high-IMS, low-EMS participants compared with both high-IMS, high-EMS participants and low-IMS participants. Blink amplitudes were not predicted to differ in response to White faces. The motivation group comparison of responses to Black faces was significant, $t(56) = 2.98$, $p < .005$, $r = .37$, indicating that blink amplitudes for high-IMS, low-EMS participants ($M = 50.65$, $SD = 4.31$) were lower than those of high-IMS, high-EMS participants and low-IMS participants ($M = 54.39$, $SD = 4.83$). The comparison did not produce a significant effect for blink responses to White faces, $t(56) = 0.45$, $p = .66$, $r = .06$ (95% CI = −.61, .77). A separate motivation group comparison examined responses to Blacks after responses to Whites had been covaried. This analysis was also significant, $t(56) = 3.08$, $p < .005$, $r = .38$, suggesting an effect of the motivation group comparison for Black faces beyond any differences in responses to White faces. Follow-up analyses of blink responses to Black faces tested each pairwise comparison between the three motivation groups. These revealed that blink amplitudes of high-IMS, low-EMS participants were significantly smaller than those of high-IMS, high-EMS participants ($M = 54.15$, $SD = 5.41$), $t(56) = 2.79$, $p < .005$, $r = .35$, as well as low-IMS participants ($M = 54.63$, $SD = 4.32$), $t(56) = 3.17$, $p < .001$, $r = .39$. A significant difference was not observed between the blink responses of high-IMS, high-EMS participants and low-IMS participants, $t(56) = 0.39$, $p = .35$, $r = .05$ (95% CI = −.65, .75). These results were consistent with the pattern of blink responses observed at the short latency, albeit with stronger effects.

The alternative hypothesis was that high-IMS, high-EMS participants would exhibit control over their initially negative response to Black faces, such that their blink amplitudes would be similar to those of high-IMS, low-EMS participants and smaller than the blinks of low-IMS participants. Although the high-IMS versus low-IMS contrast testing this prediction was significant, $t(56) = 1.75$, $p < .05$, $r = .23$, it clearly provided a poorer fit to the data than the motivation group comparison, as can be seen in Figure 2. A similar contrast comparing high- and low-IMS participants' responses to White faces was not significant, $t(56) = 1.08$, $p = .28$, $r = .14$ (95% CI = −.59, .78).

Together, these results support the hypothesis that controlled processes do not modulate the automatic affective race-biased responses of high-IMS, high-EMS participants and suggest that the race bias present in basic affective processes, as indicated by startle eyeblink response, remains active while control becomes theoretically possible. These results further suggest that long latency startle eyeblinks index automatic affective race bias, and thus corroborate the interpretation of short latency blink responses as indicating less race bias among high-IMS, low-EMS participants.

Tests of alternate explanations. To address the possibility that the motivation group differences in response to Black faces were due to an overall out-group negativity effect, we repeated the motivation group comparison analysis for blink responses to Asian faces at short and long latencies. The comparison was not significant for responses to short latency probes, $t(56) = 0.63$, $p = .53$, $r = .08$ (95% CI = −.56, .79), or to long latency probes, $t(56) = 1.29$, $p = .20$, $r = .17$ (95% CI = −.85, .85).[6] Furthermore, when responses to Asian faces were covaried from responses to Black faces, the motivation group comparisons for responses to Black faces remained significant at the 400-ms startle probe, $t(56) = 1.99$, $p < .03$, $r = .26$, and at the 4,000-ms probe, $t(56) = 2.77$, $p < .005$, $r = .35$. In combination, these analyses suggest that the individual differences in startle eyeblink responses were specific to Black faces (as opposed to racial out-groups more generally). Furthermore, this analysis suggests that differences in responses to Black faces were not due to a general unfamiliarity with Black people.

A remaining alternative explanation for the pattern of responses toward Black faces is that the motivation groups differ in basic-level affective processes. Because the pattern of startle responses to Black faces suggests the activation of negative affect, we tested this alternative explanation by conducting the motivation group comparison on startle responses to negative IAPS slides at 400 ms and 4,000 ms. This comparison was not significant at either the 400-ms probe, $t(55) = .19$, $p = .85$, $r = .03$ (95% CI = −.73, .86), or the 4,000-ms probe, $t(55) = 1.26$, $p = .21$, $r = .17$ (95% CI = −.35, .85). In addition, motivation groups did not differ significantly in response to neutral and positive stimuli at either probe latency (ts < 1.28, ps $> .21$, rs $< .13$). These analyses suggest that the pattern of affective bias toward Black faces was not due to underlying differences in general affective responses. It was notable, however, that analyses of responses to IAPS pictures replicated the well-established pattern of valence modulation in long latency blinks, $F(2, 144) = 9.34$, $p < .005$, $r = .25$. However, the pattern of attention modulation was not replicated in short latency blinks ($F < 1$). These results suggest that participants' short latency blink responses yielded a less sensitive index of affective response than our long latency blinks. This result may explain why the motivation group comparison effects for short latency blink responses were not as strong as the effects for long latency blink responses.

ATB analyses. To compare individual difference effects obtained using the IMS–EMS with those obtained using a traditional evaluative measure of prejudice, we repeated our analyses of startle eyeblinks during Black trials using ATB scores as the predictor. A median split on ATB scores was used to designate participants as falling in the high- ($M = 6.54$, $SD = .20$) versus low- ($M = 5.46$, $SD = .58$) ATB group. The median was obtained from responses to the ATB in the mass testing session ($Mdn = 6.20$, $N = 1{,}393$), a value typical of samples from this university's undergraduate population. ATB scores of high-IMS, low-EMS participants ($M = 6.43$, $SD = 0.43$) and high-IMS, high-EMS participants ($M = 6.13$, $SD = 0.46$) were both greater than those of low-IMS participants ($M = 5.45$, $SD = 0.75$, ts > 3.28, ps $< .005$).

---

[6] We conducted 2 (motivation group comparison) × 2 (race of face) ANOVAs on participants' responses to Black and Asian faces at 400-ms and 4,000-ms startle probe latencies. At 400 ms, the ANOVA produced a main effect for race of face, $F(1, 56) = 14.57$, $p < .001$, and a marginal effect for motivation group comparison, $F(1, 56) = 2.86$, $p = .10$, but no interaction ($F < 1$). At 4,000 ms, the ANOVA produced main effects for race of face, $F(1, 56) = 17.05$, $p < .001$, and for motivation group comparison, $F(1, 56) = 7.60$, $p = .10$, but no interaction ($F < 1$). The patterns of these effects were similar and suggested a higher level of race bias toward Black faces relative to Asian faces. In addition, high-IMS, low-EMS participants exhibited less blink modulation to both Black faces and Asian faces compared with other participant groups.

To test the ability of ATB scores to account for patterns of variance in blink responses, a 2 (ATB: high vs. low) × 2 (race of face: Black vs. White) mixed-factorial ANOVA was conducted for blink responses at each latency. Analysis of short latency blink responses revealed a marginal main effect of ATB, $F(1, 53) = 3.03$, $p = .09$, $r = .23$, such that the blinks of high-ATB participants ($M = 47.56$, $SD = 3.91$) were larger than those of low-ATB participants ($M = 46.28$, $SD = 3.19$). Effects did not reach significance for race of face ($F < 1$), or the interaction, $F(1, 53) = 1.35$, $p = .25$, $r = .16$. Analysis of long latency blinks produced an effect for race of face, $F(1, 53) = 12.35$, $p < .001$, $r = .43$, such that blinks were larger to Black faces ($M = 52.84$, $SD = 5.06$) than to White faces ($M = 50.03$, $SD = 3.86$). Effects for ATB and the interaction did not reach significance (Fs < 1). These results were consistent with Devine et al.'s (2002) finding that traditional attitude measures of prejudice were unable account for variability in levels of implicit race bias.

We next examined the effect of ATB separately for short and long latency blink responses to Black faces. ATB did not predict blink responses at either 400 ms, $t(53) = .50$, $p = .62$, $r = .07$ (95% CI = −.60, .78), or 4,000 ms, $t(53) = .96$, $p = .34$, $r = .13$ (95% CI = −.46, .82).[7] Because participants were not selected on the basis of extreme ATB scores, it was appropriate to examine ATB as a continuous variable. We therefore conducted supplementary analyses using continuous ATB and eyeblink scores. Correlations of participants' ATB scores with residual blinks to Black faces, after they had been regressed on blink responses to White faces, were not significant at either 400 ms, $r(58) = -.05$, $p = .71$, or at 4,000 ms, $r(58) = -.13$, $p = .36$. Finally, to provide a strong test of our a priori comparison, we repeated the motivation group comparison for short and long latency blink responses to Black faces while including ATB as a covariate in analyses of covariance (ANCOVAs). The ANCOVAs produced results similar to the planned comparisons reported above, such that high-IMS, low-EMS participants exhibited less blink modulation than other participants at the 400-ms latency, $F(1, 52) = 3.39$, $p < .04$, $r = .25$, and at the 4,000-ms latency, $F(1, 52) = 9.40$, $p < .005$, $r = .39$. Neither ANCOVA produced an effect for ATB (Fs < 1). Our results demonstrate that ATB scores were not directly associated with eyeblink responses to Black faces and that, when covaried from analysis, they did not alter the predictive effects of IMS–EMS. It should be noted that our participant-selection criteria were designed to enhance our power in detecting IMS–EMS motivation-group effects; participants were not selected for extreme ATB scores. This should be taken into consideration when comparing the null effects of ATB with the stronger effects of IMS–EMS motivation group.

Discussion

The results of Study 1 provide evidence for individual differences in automatic affective bias toward Black people. Across startle probe latencies, high-IMS, low-EMS participants exhibited lower levels of automatic affective bias than other participants. This finding is consistent with Devine et al.'s (2002) proposal that low levels of race bias among high-internal, low-external individuals are a result of their having internalized their nonprejudiced beliefs to the point that their prejudice-related responses are effectively regulated across situations. The blink responses of high-IMS, high-EMS participants, on the other hand, reflected relatively higher levels of affective race bias that were inconsistent with their nonprejudiced beliefs. Indeed, Devine (1989) characterized individuals showing a dissociation of automatic and controlled prejudiced responses as egalitarians who have not yet broken the prejudice habit. Similar dissociations have been reported for high-IMS, high-EMS participants in recent work by Devine et al. (2002) and Plant and Devine (2002). These results support the assertion that some low-prejudice individuals experience spontaneous and unintentional negative feelings toward Black people, whereas other low-prejudice individuals do not (Devine et al., 1991). Lastly, the blink responses of low-IMS participants were consistent with their high level of self-reported prejudice. By replicating previous patterns of implicit race bias using a physiological measure of basic affect, the results of Study 1 suggest that affective processes are present in the automatic activation of race bias.

An intriguing result of Study 1 was that the pattern of blink responses observed at the short latency was also observed at the long latency, when controlled processes were theoretically available. A comparison between the short and long latency effect sizes for motivation group responses to Black faces was not significant, $t(56) = 1.05$, $p = .30$, $r = .17$ (95% CI = −.15, .48). Taking into account that CIs for startle eyeblink differences are typically inflated, this analysis suggests that blinks at short and long latencies indexed the same underlying construct, namely, automatic affective race bias. This conclusion is further supported by significant correlations between participants' short and long latency blink responses to Black faces, $r(56) = -.42$, $p < .002$, and between short and long latency residualized blink responses to Black faces, in which White faces were covaried, $r(56) = -.38$, $p < .004$. Thus, it is not surprising that we did not observe control in the startle eyeblink responses of high-IMS, high-EMS participants. Although high-IMS, high-EMS participants have typically reported low levels of prejudice on explicit measures (Devine et al., 2002; Plant & Devine, 1998; Plant et al., 2003) and have not discriminated by race in certain overt behaviors (Amodio & Devine, 2000), they have been shown to exhibit relatively high levels of implicit bias (Devine et al., 2002). Thus, our conclusion that both short and long latency blink responses to Black faces reflected automatic levels of affective race bias is consistent with past research. Nevertheless, questions remain regarding how control over initially biased affect may be implemented, particularly for high-IMS, high-EMS individuals.

Guglielmi (1999) recently suggested that whereas basic-level affective responses are more difficult to control, belief-based pro-

---

[7] This finding, as well as Phelps et al.'s (2000) finding, is inconsistent with some previous work suggesting that individual differences in racial attitudes moderate physiological responses to race-relevant stimuli (e.g., Vanman et al., 1997). Recent research (e.g., Devine et al., 2002; Plant & Devine, 1998) has indicated that there is variability among low-prejudice people that is not captured by traditional measures of evaluation such as the MRS (used by Vanman et al., 1997, and Phelps et al., 2000) and the ATB (used in the present research). One possible explanation for this inconsistency among past studies is that in any given sample, the low-prejudice participants could be composed of mostly high-IMS, low-EMS participants or high-IMS, high-EMS participants. Depending on the composition of the low-prejudice group, very different outcomes could be observed, as demonstrated in the current article.

cesses may operate more effectively in other response channels such as self-reports. In the context of prejudice research, Guglielmi proposed a framework whereby automatic (uncontrolled) levels of affective processing are appropriately measured with unobtrusive physiological indices, and controlled (explicit) levels of affective processing are more sensitively captured by self-reports. Examples of affective control occurring at the self-report, but not physiological, level have been documented in the psychophysiological literature (Schwartz & Kline, 1995, Weinberger, Schwartz, & Davidson, 1979). Past theorizing in the social psychological literature has suggested that successful control requires (a) an awareness that bias is possible and (b) the availability of cognitive resources to implement controlled responses (Wegner & Bargh, 1998). Because the procedure used in Study 1 minimized participants' awareness that affective responses might be biased and thus provided little opportunity for deliberative control, we conducted a second study that used a procedure more amenable to controlled processing.

Study 2

In Study 2, we tested the hypothesis that controlled affective responses would be exhibited by high-IMS, high-EMS participants when made via confidential self-reports. The self-report format permitted participants to be aware that their affect was being measured and that it could be biased. It also allowed participants sufficient cognitive resources for control to be implemented. Using a design analogous to that of Study 1, participants viewed pictures of Black, White, and Asian faces and made ratings of their subjective experience of pleasantness and arousal, the two dimensions theorized to underlie basic affective responses (Feldman Barrett & Russell, 1998; M. K. Greenwald, Cook, & Lang, 1989; Lang et al., 1990; Osgood, 1952). Our primary hypothesis, derived from Guglielmi's (1999) framework, was that controlled, belief-based responses to Black faces would be evident in self-reports of affect. We predicted that high-IMS participants (irrespective of EMS score) would report higher levels of pleasantness in response to Black faces than low-IMS participants, and that the ratings of high-IMS, high-EMS participants would not differ from those of high-IMS, low-EMS participants. That is, whereas high-IMS, high-EMS participants in Study 1 responded with higher levels of affective bias toward Black faces on the startle eyeblink measure, this bias should not be evident on self-report measures that are more sensitive to controlled, belief-based responding.

Method

Participants and design. Ninety White students (28 men, 62 women) enrolled in Introductory Psychology participated voluntarily for extra course credit. As in Study 1, the IMS ($\alpha = .80$) and EMS ($\alpha = .82$) were completed during a mass testing session held early in the semester. Participants were selected to represent levels of IMS and EMS matching the extreme scores of Study 1 participants, although experimenters were blind to participants' scores. As in Study 1, participants represented the high-IMS, low-EMS group ($M_{\text{IMS}} = 8.76$, $M_{\text{EMS}} = 2.62$), the high-IMS, high-EMS group ($M_{\text{IMS}} = 8.75$, $M_{\text{EMS}} = 7.22$), and the low-IMS group ($M_{\text{IMS}} = 6.47$, $M_{\text{EMS}} = 4.58$). The experimental design was a 3 (motivation group: high IMS, low EMS vs. high IMS, high EMS vs. low IMS) × 2 (race of face: Black vs. White) mixed factorial, with repeated measures on race of face. The dependent measures were self-reported affect and arousal.

Procedure and materials. On entering the laboratory, groups of up to 3 participants were seated approximately 6 ft (1.83 m) in front of a 17-inch (43.18 cm) computer monitor. Large dividers separated participants into private individual carrels. The lighting in the room was dimmed to enhance the viewing of images presented on the monitor. Participants were told they would rate each picture according to how pleasant and excited they felt personally in response to each picture, and that all responses would be confidential. They were specifically instructed not to make identifying marks on the response forms and were told that they would place completed forms in a slotted drop box along with those of the other participants to further ensure their confidentiality. In past research, this procedure has been shown to successfully prevent the engagement of participants' external concerns (Devine et al., 2002; Plant & Devine, 1998).

Ratings for valence and arousal were made separately on scales ranging from 1 (*unpleasant*) to 9 (*pleasant*), and 1 (*calm*) to 9 (*excited*), respectively. Research based on theorizing dating back to Wundt (1896) and Osgood (1952) has shown that the dimensions of valence and arousal underlie affective responses (e.g., Bradley & Lang, 1994; Lang, Bradley, & Cuthbert, 1992, 1997; Lang et al., 1990; Lang, Greenwald, Bradley, & Hamm, 1993; Vrana, Spence, & Lang, 1988). To increase participants' understanding of these basic affective constructs, the endpoints of these scales were labeled on the coversheet with synonyms used in much emotion research on the startle eyeblink response (Lang et al., 1997). For the valence dimension, these included *annoyed*, *unhappy*, and *unsatisfied* at the "unpleasant" endpoint and *pleased*, *happy*, and *content* at the "pleasant" endpoint. For the arousal dimension, these included *relaxed*, *sluggish*, and *sleepy* at the "calm" endpoint and *stimulated*, *wide-awake*, and *jittery* at the "excited" endpoint. Thus, the scales measured very general experiences of pleasure and arousal. This approach was advantageous because it allowed for quick, gut-level reports of general affect.

Participants were then shown the same series of faces used in Study 1, in the same quasi-random order. Each picture was presented for 6 s and was followed by a 5-s ITI during which participants made their ratings. On completion, participants placed their rating forms into the drop box. Participants were then debriefed, thanked, and dismissed.

Results

Ratings of pleasantness and arousal were averaged for each participant within the race of face condition. Pleasantness and arousal indices demonstrated high reliability ($\alpha\text{s} > .87$), and Cronbach's alphas were comparable across conditions. One participant failed to complete the pleasantness ratings and was not included in analyses involving this variable. In a preliminary analysis, participant sex did not moderate the effect of Motivation Group × Race of Face on affect or arousal ratings ($F\text{s} < 1$) and was excluded from further analyses.

Reported affective responses. On the basis of Guglielmi's (1999) theorizing, we hypothesized that high-IMS, high-EMS participants, who exhibited automatic race bias in their startle eyeblink responses in Study 1, would not reveal race bias in their self-reports. Thus, we predicted that high-IMS participants would report more positive ratings of Black faces compared with low-IMS participants but that these groups would not differ in their ratings of White faces. Hence, self-reported affective responses were analyzed using a 2 (IMS: high vs. low) × 2 (race of face: Black vs. White) mixed-factorial ANOVA, with race of face as the repeated measure. This analysis produced a main effect for IMS, $F(1, 87) = 4.74$, $p < .04$, $r = .23$, and no effect for race of face ($F < 1$). This main effect was qualified by a marginally significant interaction, $F(1, 87) = 3.14$, $p = .08$, $r = .19$. Mean levels self-reported affect for each motivation group are illustrated in Figure 3.

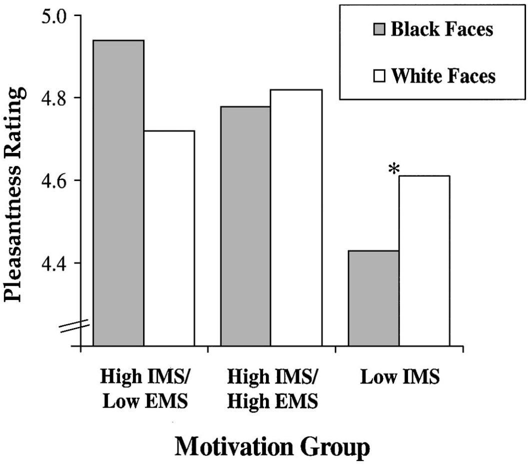

Figure 3. Mean levels of self-reported affective valence in response to Black and White faces as a function of motivation group. Responses ranged from 1 (*unpleasant*) to 9 (*pleasant*). An asterisk denotes a significant difference ($p < .05$) in Black versus White responses within motivation group. Between-subjects differences are reported in the text. IMS = Internal Motivation Scale; EMS = External Motivation Scale.

Separate IMS comparisons examined our a priori predictions for self-reported affective responses to Black faces and White faces separately. The IMS comparison of pleasantness ratings of Black faces was significant, $t(87) = 2.65$, $p < .005$, $r = .27$, such that ratings made by high-IMS participants ($M = 4.86$, $SD = 0.61$) were more positive than those made by low-IMS participants ($M = 4.43$, $SD = 0.94$).[8] Additional analyses demonstrated that ratings of Black faces did not differ significantly between high-IMS, low-EMS participants ($M = 4.94$, $SD = 0.60$) and high-IMS, high-EMS participants ($M = 4.78$, $SD = 0.61$), $t(86) = 0.80$, $p = .43$, $r = .09$ (95% CI = −.45, .83). However, the ratings of low-IMS participants were significantly lower than both high-IMS, low-EMS participants, $t(86) = 2.67$, $p < .01$, $r = .29$, and high-IMS, high-EMS participants, $t(86) = 2.04$, $p < .04$, $r = .25$. The IMS comparison did not predict ratings of pleasantness in response to White faces, $t(87) = 1.05$, $p = .30$, $r = .11$ (95% CI = −.42, .88).

To rule out the alternative explanation that IMS differences in responses to Black faces were due to the general out-group negativity, it was important to show that the IMS did not also predict responses to Asian faces. To this end, the IMS comparison was conducted for responses to Asian faces and was found to be marginal, $t(87) = 1.75$, $p = .08$, $r = .18$ (95% CI = −.12, .88).[9] Asian faces were judged somewhat more positively by high-IMS participants ($M = 4.54$, $SD = .78$) compared with low-IMS participants ($M = 4.19$, $SD = 1.06$). An additional analysis showed that IMS continued to predict responses to Black faces after responses to Asian faces had been covaried, $t(87) = 1.95$, $p < .03$, $r = .20$. Together, these analyses cast doubt on the possibility that the IMS effect for Black faces was due to a general out-group negativity or to the relative novelty of Black faces among our White participants.

Arousal ratings. Four participants failed to complete the arousal portion of ratings correctly, and their scores were excluded from analyses of arousal. A preliminary 2 (IMS: high vs. low) × 2 (race of face: Black vs. White) mixed-factorial ANOVA, with race of face as the repeated measure, did not produce significant effects for race of face ($F < 1$); for IMS, $F(1, 83) = 1.86$, $p = .18$, $r = .15$; or for the interaction ($F < 1$). The arousal ratings of high- and low-IMS participants did not differ in response to Black faces, $t(83) = 1.35$, $p = .18$, $r = .12$ (95% CI = −.30, .85), or to White faces, $t(83) = 1.09$, $p = .28$, $r = .14$ (95% CI = −.41, .83). Additionally, arousal ratings did not differ in response to Asian faces, $t(83) = .45$, $p = .65$, $r = .05$ (95% CI = −.61, .77), although the large CIs prevent strong conclusions that a true difference does not exist between high and low internally motivated individuals.

In light of these null effects, we considered the possibility that participants were simply not aware of the type of subtle, rapid changes in physiological arousal detected by the startle eyeblink measure used in Study 1. That is, whereas people can detect a visceral arousal response when they view extremely negative (e.g., accident victims) or extremely positive (e.g., opposite-sex nudes) images (e.g., Lang et al., 1997), it is much more difficult to report arousal changes in response to the relatively neutral images of faces with neutral expressions. On the basis of past theorizing and research that valence and arousal are independent dimensions of affect (Bradley et al., 1993; Lang et al., 1990; Russell, 1980), we would expect that to the extent that participants' self-reports were sensitive to affect-related changes in arousal, reports of pleasantness and arousal should be uncorrelated. An examination of these correlations revealed that pleasantness and arousal were in fact correlated in ratings of Black faces ($r = .35$, $p < .01$), White faces ($r = .25$, $p = .05$), and Asian faces ($r = .30$, $p < .02$), indicating that higher arousal was associated with more pleasantness. One explanation for these correlations is that despite the instructions, some participants interpreted the measure as pertaining to sexual arousal or interpersonal attraction. An alternative explanation is that participants associated the synonyms provided for arousal on the questionnaire with valence, such that "aroused" words seemed slightly more positive than "unaroused" words in the context of viewing neutral faces. Overall, the pattern of self-reported arousal did not correspond to the well-established pattern found in the emotion-modulated startle literature (whereby arousal and valence should be uncorrelated), and suggests that self-reported arousal did not assess the type of physiological arousal that underlies modulation of the startle response.

---

[8] Because only women participated in Study 1, we conducted the primary Study 2 analyses using the responses of female participants only. A test of the planned IMS comparison was significant in response to Black faces, $t(59) = 2.16$, $p < .02$, $r = .27$, such that the ratings of high-IMS participants were more positive ($M = 4.93$, $SD = 0.67$) than the ratings of low-IMS participants ($M = 4.48$, $SD = 0.92$). The comparison was not significant for responses to White faces, $t(59) = 1.25$, $p = .22$, $r = .16$, or to Asian faces, $t(59) = 1.62$, $p = .11$, $r = 21$.

[9] A 2 (IMS) × 2 (race of face) mixed-factorial ANOVA for self-reported affective responses to Black and Asian faces produced significant effects for the IMS, $F(1, 87) = 5.95$, $p < .02$, and for race of face, $F(1, 87) = 11.14$, $p < .001$, but there was not an interaction ($F < 1$). The main effects suggested that ratings were generally more positive for Asian than for Black faces, and that high-IMS participants had a general tendency to rate faces more positively than low-IMS participants.

Discussion

The results of Study 2 suggest that controlled, belief-based processes are implemented in more deliberative affective responses to Black faces. Whereas high-IMS, high-EMS participants exhibited high levels of race bias to both short and long latency startle probes in Study 1, their self-reported affective responses revealed low levels of bias in Study 2. Our findings are consistent with Guglielmi's (1999) proposal that self-reports provide a more appropriate and sensitive measure of controlled processes than unobtrusive physiological indices of basic affective processes. As in Study 1, it should be noted that the nonsignificant results we reported were accompanied by large CIs. These large CIs suggest that relatively high levels of variability contributed to the null effects revealed in these analyses, and therefore these findings should be interpreted with some caution. Nevertheless, it is impressive that photographs of White and Black faces with neutral expressions were able to cause differences in high- and low-IMS participants' self-reported affective responses.

One may wonder if high-IMS, high-EMS participants were deceptive in providing low levels of race bias in self-reported responses. Converging pieces of evidence suggest that this possibility is unlikely. First, we went to great lengths to emphasize participants' confidentiality in order to mitigate social desirability concerns, thereby reducing the motivation to intentionally misreport their responses. Second, past research has found that high-IMS, high-EMS participants typically respond without prejudice in their self-reported beliefs and attitudes as well as in their deliberative behaviors in private and in public, suggesting that self-reports of high-IMS, high-EMS participants were veridical representations of their subjective affective experiences (Amodio & Devine, 2000; Plant & Devine, 1998; Plant et al., 2003). In light of this past work, we can be fairly confident in our conclusion that for some individuals (e.g., high-IMS, high-EMS participants), controlled, nonprejudiced responses are effectively conveyed in self-reports despite the activation of automatic affective racial bias.

Comparison of Effect Sizes Across Studies

We conducted cross-study comparisons to examine the effects of motivation group on automatic affective responses in Study 1 and controlled affective responses in Study 2. Specifically, we conducted these analyses to assess whether high-IMS, low-EMS participants and high-IMS, high-EMS participants differed in their automatic but not in their controlled affective responses to Black faces. Cross-study comparisons of effect sizes revealed that the difference in long latency startle eyeblink responses was significantly larger than the difference in self-reported affect ($z = 1.78$, $p < .04$). The difference in short latency startle eyeblink responses was also larger than the difference in self-reported affect, but not at a significant level ($z = .91$, $p = .18$, 97.5% CI $= -0.73$, 2.55). Although interpretations regarding the findings of short latency blink differences must be interpreted more cautiously, the findings for long latency blink differences were clear. Overall, these analyses supported our conclusion that high-IMS, high-EMS participants exhibited race bias in their automatic but not in their controlled affective responses, relative to high-IMS, low-EMS participants.

General Discussion

Taken together, the present results demonstrate that there are predictable individual differences in basic level affective processes associated with the automatic activation of race bias. They further suggest that controlled, belief-based responses are implemented not at the basic affective level but in more deliberative channels of affective expression. By linking race bias directly to basic affective processes, the present research provided support for the widely acknowledged but seldom tested assumption that affective responses represent a major component of racial bias (Allport, 1954; Hamilton & Mackie, 1993; Mackie & Smith, 1998; Smith, 1993). Whereas much recent research on implicit race bias has focused on theoretically "cognitive" constructs, such as the spreading activation of evaluative or semantic associations, future research will benefit from a broader conceptualization of this process that includes affective processes (Hamilton, 1981). In addition, identifying the affective component of race bias is important because there are known psychophysiological mechanisms associated with it. By associating automatic affective race bias with patterns of startle eyeblink response, we can link the activation of race bias to basic mechanisms of emotional learning (e.g., involving the amygdala). Through this link, researchers may benefit from the vast neuroscience literature on the functions and neuroanatomy of these mechanisms to learn more about how this bias may be acquired and perhaps diminished. We address the potential of this approach in more detail in *Emotional Learning, Neural Plasticity, and Attitude Change*, below.

Automatic Affective Race Bias

The primary contribution of Study 1 was to identify individual differences in implicit affective race bias. A major strength of the startle eyeblink assessment of automatic affective activation is that it does not require a behavioral response from the participant and therefore allows for a more precise temporal measurement of rapidly unfolding implicit processes. The measurement of blink responses provides a sensitive index of affect activation and, by allowing for assessments of affect at very early stages of processing, permits a strong test of the hypothesis that individuals differ in their levels of automatic race bias. By comparison, the evidence for implicit bias provided by reaction-time measures can be ambiguous because the time frame in which responses are made is not well controlled (Payne, 2001). That is, responses on such measures can vary in reaction time from as little as 250 ms to several seconds, and hence the degree to which any given reaction time is affected by controlled processes is unclear. Nevertheless, by replicating previously observed individual differences in implicit bias with startle modulation methodology, the present results bolster our confidence in the implications of internal and external motivations to respond without prejudice for automatic responses. Finally, the patterns of eyeblink responses obtained in Study 1 suggest an affective component of implicit race bias. Theorizing regarding the interplay between affective and cognitive processes suggests that they are tightly interrelated at the implicit level (Kihlstrom, Mulvaney, Tobias, & Tobis, 2000), and further suggests the possibility that implicit affective processes may precede and guide cognition (Zajonc, 1980). Because the present study represents one of the first investigations of implicit affective race bias, much future research will be needed to delineate the interplay of implicit affect and cognition in the context of race bias.

Control of Affective Race Bias

Evidence of controlled, belief-consistent responses was not observed in the startle eyeblink responses of Study 1 participants but was evident in the self-reported affective responses in Study 2. Of the three motivation groups examined in these studies, only participants in the high-IMS, high-EMS group were expected to exhibit a dissociation in their automatic and controlled responses. These participants showed no physiological evidence of control over the activation of race-related negative affect, exhibiting a similar level of negative affect to Black faces at long probe latencies as did low-IMS participants (Study 1). In contrast, their self-reported responses to Black faces were more positive than those of low-IMS participants and did not differ significantly from the responses of high-IMS, low-EMS participants (Study 2; however, the lack of a statistically significant difference must be qualified by a relatively large CI). Although this finding is consistent with Guglielmi's (1999) framework, the degree to which basic-level affective bias may be modified by control efforts is unclear. Indeed, the possibility that basic forms of race-related negative affect cannot be controlled once activated would have troubling implications for high-IMS, high-EMS individuals, because it may have serious and potentially pernicious unintended consequences for judgments and behavior (e.g., Zajonc, 1998). Recent theorizing has addressed the varying levels of control that may exist across response channels (Devine & Monteith, 1999; Rudman, Ashmore, & Gary, 2001), and the present findings further suggest that the effectiveness of control efforts may differ in affective and cognitive processes. Although the present research examined affective responses at the extremes of automaticity and control, future research is needed to explore the relative contributions of automatic and controlled processes across a fuller range of affective responses.

Tests of Alternative Explanations in Research on Race Bias

We made a special effort to consider alternative explanations for the findings we obtained. Across studies, we found support for our hypothesis that individual differences in internal and external motivations to respond without prejudice toward Black people were specific to Black faces and not to racial out-groups in general. We also assessed whether individual differences in responses to Black faces were not merely due to White participants' unfamiliarity with Black people. We found that the individual difference measures did not predict responses to a different but similarly unfamiliar minority group, Asians. Testing this explanation is particularly important when examining rapidly executed responses such as short latency startle eyeblinks and short stimulus onset asynchrony reaction times. These rapidly activated processes can be particularly sensitive to contextual inconsistencies, such as the relative novelty of a stimulus, in addition to purely evaluative or semantic features (Bradley et al., 1999; Brendl, Markman, & Messner, 2001; Whalen, 1998). One possible explanation for why we did not obtain group differences in startle eyeblink responses to Asian faces is that stereotypes of Asians do not include threatening attributes (Niemann, Jennings, Rozelle, Baxter, & Sullivan, 1991). Indeed, it is possible that implicit bias toward Asians might be driven predominantly by stereotypes and not by basic-level negative affect and that race bias associated with negative affect may lead to different forms of discrimination than race bias not associated with negative affect. These possibilities highlight interesting avenues for future research. Finally, because our studies focused on individual differences in affective responses associated with race bias, it was also important to show that the individual difference groups did not differ in their general affective responses. In this regard, our results did not indicate a relationship between participants' internal and external motivations to respond without prejudice and their startle eyeblink responses to IAPS pictures.

Differentiation of Affective Responses to Out-Groups

The present analysis examined the most basic level of affective responses along the continua of valence and arousal, as theorized by several scientists (e.g., Lang et al., 1990; Osgood, 1952; Wundt, 1896). This theoretical framework is most amenable to the analysis of rapidly occurring affective responses associated with subcortical neural structures such as the amygdaloid complex. However, affective responses to out-groups can be elaborated and differentiated through a process of appraisal to represent specific relationships between perceivers and the targets of prejudice (Mackie & Smith, 1998; Smith, 1993). Our intent in the present report was to examine the affective starting point from which more complex emotional and cognitive processes may develop. Exploring the interplay between the basic-level affect and the more elaborated approaches provides fertile ground for future work into the physiology and subjective experience of reactions to out-groups.

Emotional Learning, Neural Plasticity, and Attitude Change

In light of having detected individual differences in implicit levels of race-biased affect, an important new question concerns the malleability of implicit affective bias. This question is central to efforts aimed at prejudice reduction. Devine et al. (2002) have proposed a model whereby individuals may come to show less prejudice over time through a process of internalization. That is, many individuals endorse egalitarianism yet sometimes react negatively toward Blacks. This characterization describes the high-IMS, high-EMS participants. By associating implicit race bias with amygdala activity (via the startle response), we can draw from the neuroscience literature pertaining to affective plasticity to better understand the challenges of prejudice reduction. For example, the amygdala is known to encode well-learned affective associations that are not easily altered (Rolls, 1999). The work of Rolls (2000) has demonstrated that although changes in the prefrontal cortex in response to learning contingencies occur rapidly, changes at the amygdalar level require significantly more reinforcement. Thus, such associations are more resistant to change. It appears that the difficulties researchers often encounter when attempting to change prejudiced attitudes may have roots in the nature of neural mechanisms of emotional learning. Furthermore, given that the amygdala responds more quickly to stimuli than do structures in the prefrontal cortex (LeDoux, 1996), Rolls's work suggests that until a response is unlearned at the more basic levels, effortful controlled processes must intervene to prevent unwanted behavioral consequences. Devine et al. (2002) have proposed a similar process, albeit at a different level of analysis, for the unlearning of racial prejudice, whereby reducing prejudice in one's responses involves the unlearning of well-elaborated negative associations formed in the context of a culture in which they were reinforced. Integrating the work of Rolls and Devine et al., we suggest that the nonprejudiced self-reported responses made by high-IMS, high-EMS individuals may be associated with activity of the prefrontal cortex but not subcortical structures. As their nonprejudiced beliefs become more internalized, these individuals would be expected to rely less on prefrontal processing when responding without prejudice because race bias would be less activated at lower levels in subcortical circuitry. Findings from the animal literature suggest that basic-level changes in behavioral associations are a lengthy and arduous process (e.g., Rolls, 1999), and therefore individuals in the process of "breaking the prejudice habit" can only achieve their egalitarian goals through practice over time (Devine & Monteith, 1993; Monteith, 1993; Monteith, Ashburn-Nardo, Voils, & Czopp, 2002). Although the present findings offer support for this idea from a cross-sectional sample, future studies of startle eyeblink differences using longitudinal designs would provide a more compelling test of how changes in affective responses, indexed by startle eyeblinks, correspond to the internalization of nonprejudiced beliefs and changes in implicit race bias level.

In a broader sense, our results demonstrate the fruitful potential that lies in examining social psychological phenomena at the intersection of motivational, affective, cognitive, and neuroscientific approaches (Cacioppo, Berntson, Sheridan, & McClintock, 2000; Ito & Cacioppo, 2001; Ochsner & Lieberman, 2001). Although the research literatures of each of these disciplines are expansive and diverse, we believe that the synthesis of these approaches will yield a more comprehensive analysis of many psychological phenomena. At the same time, the findings of the present research reaffirm the utility of more traditional methodological formats (e.g., self-reports) that continue to be well-suited to assessments of consciously held beliefs and subjective experiences. At a time when "indirect" measures, such as reaction time and physiological indices, are having a large impact on social psychological research, the present work attests to the importance of considering multiple levels of response in understanding psychological processes.

References

Adolphs, R., Tranel, D., & Damasio, A. R. (1998, June 4). The human amygdala in social judgment. *Nature, 393,* 470–474.

Alloy, L. B, Abramson, L. Y., Raniere, D., & Dyller, I. M. (1999). Research methods in adult psychopathology. In P. C. Kendall, J. N. Butcher, & G. N. Holmbeck (Eds.), *Handbook of research methods in clinical psychology.* New York: Wiley.

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

Amodio, D. M., & Devine, P. G. (2000, January). *Strategic concealment of race bias.* Poster session presented at the first annual meeting of the Society for Personality and Social Psychology, Nashville, TN.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1993). Pictures as prepulse: Attention and emotion in startle modification. *Psychophysiology, 30,* 541–545.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1999). Affect and the startle reflex. In M. E. Dawson, A. M. Schell, & A. H. Böhmelt (Eds.), *Startle modification: Implications for neuroscience, cognitive science, and clinical science.* New York: Cambridge University Press.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25,* 49–59.

Brendl, M. C., Markman, A. B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology, 81,* 760–773.

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

Buswell, B. N., & Devine, P. G. (2000, February). *Internal and external motivations to respond without prejudice to fat people.* Poster session presented at the first annual meeting of the Society for Personality and Social Psychology, Nashville, TN.

Cacioppo, J. T., Berntson, G. G., Sheridan, J. F., & McClintock, M. K. (2000). Multilevel integrative analyses of human behavior: Social neuroscience and the complementing nature of social and biological approaches. *Psychological Bulletin, 126,* 829–843.

Center for the Study of Emotion and Attention. (1995). *The international affective picture system* [digitized computer images]. Gainesville: Center for Research in Psychophysiology, University of Florida.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24,* 349–354.

Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & Friesen, W. V. (1990). Approach–withdrawal and cerebral asymmetry: Emotional expression and brain physiology. I. *Journal of Personality and Social Psychology, 58,* 330–341.

Davidson, R. J., & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Science, 3,* 11–21.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11,* 227–268.

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

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

Devine, P. G., & Monteith, M. J. (1999). Automaticity and control in stereotyping. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 339–360). New York: Guilford Press.

Devine, P. G., Monteith, M., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology, 60,* 817–830.

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.

Dikman, Z. V., & Allen, J. J. B. (2000). Error monitoring during reward and avoidance learning in high- and low-socialized individuals. *Psychophysiology, 37,* 43–54.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology. II. *Journal of Personality and Social Psychology, 58,* 342–353.

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

Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology, 74,* 967–984.

Filion, D. L., Dawson, M. E., & Schell, A. M. (1998). The psychological significance of human startle eyeblink modification: A review. *Biological Psychology, 47,* 1–43.

Forster, K. I., & Forster, J. C. (in press). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers.*

Globisch, J., Hamm, A. O., Esteves, F., & Öhman, A. (1999). Fear appears fast: Temporal course of startle reflex potentiation in animal fearful subjects. *Psychophysiology, 36,* 66–75.

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

Greenwald, M. K., Cook, E. W., & Lang, P. J. (1989). Affective judgment and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. *Journal of Psychophysiology, 3,* 51–64.

Grillon, C., & Davis, M. (1995). Acoustic startle and anticipatory anxiety in humans: Effects of monaural right and left ear stimulation. *Psychophysiology, 32,* 155–161.

Grolnick, W. S., & Ryan, R. M. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology, 52,* 890–898.

Guglielmi, R. S. (1999). Psychophysiological assessment of prejudice: Past research, current status, and future directions. *Personality and Social Psychology Review, 3,* 123–157.

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

Hamilton, D. L., & Mackie, D. M. (1993). Cognitive and affective processes in intergroup perception: The developing interface. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in group perception* (pp. 1–12). San Diego, CA: Academic Press.

Harmon-Jones, E., & Allen, J. J. B. (1997). Behavioral activation sensitivity and resting frontal EEG asymmetry: Covariation of putative indicators related to risk for mood disorders. *Journal of Abnormal Psychology, 106,* 159–163.

Harmon-Jones, E., & Allen, J. J. B. (1998). Anger and prefrontal brain activity: EEG asymmetry consistent with approach motivation despite negative affective valence. *Journal of Personality and Social Psychology, 74,* 1310–1316.

Harmon-Jones, E., & Allen, J. J. B. (2001). The role of affect in the mere exposure effect: Evidence from psychophysiological and individual differences approaches. *Personality and Social Psychology Bulletin, 27,* 889–898.

Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *NeuroReport, 11,* 2351–2355.

Hawk, L. W., & Cook, E. W. (2000). Independence of valence modulation and prepulse inhibition of startle. *Psychophysiology, 37,* 5–13.

Hayes, W. L. (1988). *Statistics* (4th ed.). Fort Worth, TX: Holt, Rinehart, & Winston.

Ito, T. A., & Cacioppo, J. T. (2001). Affect and attitudes: A social neuroscience approach. In J. P. Forgas (Ed.), *Handbook of affect and social cognition* (pp. 50–74). Mahwah, NJ: Erlbaum.

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

Kihlstrom, J. F., Mulvaney, S., Tobias, B. A., & Tobis, I. P. (2000). The emotional unconscious. In E. Eich, J. F. Kihlstrom, G. H. Bower, J. P. Forgas, & P. M. Niedenthal (Eds.), *Cognition and emotion* (pp. 30–86). New York: Oxford University Press.

Klonis, S., & Devine, P. G. (2000, February). *Internal and external motivations to respond without sexism.* Poster session presented at the first annual meeting of the Society for Personality and Social Psychology, Nashville, TN.

Koestner, R., Bernieri, F., & Zukerman, M. (1992). Self-regulation and consistency between traits, attitudes, and behavior. *Personality and Social Psychology Bulletin, 18,* 52–59.

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

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1992). A motivational analysis of emotion: Reflexˆcortex connections. *Psychological Science, 3,* 44–49.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system (IAPS): Technical manual and affective ratings* (Tech. Rep. No. A-4). Gainesville: The Center for Research in Psychophysiology, University of Florida.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30,* 261–273.

Larson, C. L., Ruffalo, D., Nietert, J. Y., & Davidson, R. J. (2000). Temporal stability of the emotion-modulated startle response. *Psychophysiology, 37,* 92–102.

Leary, M. R. (1983). Social anxiousness: The construct and and its assessment. *Journal of Personality Assessment, 47,* 66–75.

LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life.* New York: Simon & Schuster.

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

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

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

Monteith, M. J. (1993). Self-regulation of stereotypical responses: Implications for progress in prejudice reduction. *Journal of Personality and Social Psychology, 65,* 469–485.

Monteith, M. J., Ashburn-Nardo, L., Voils, C. I., & Czopp, A. M. (2002). Putting the brakes on prejudice: On the development and operation of cues for control. *Journal of Personality and Social Psychology, 83,* 1029–1050.

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

Niemann, Y. F., Jennings, L., Rozelle, R. M., Baxter, J. C., & Sullivan, E. (1991). Use of free responses and cluster analysis to determine stereotypes of eight groups. *Personality and Social Psychology Bulletin, 20,* 379–390.

Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *American Psychologist, 56,* 717–734.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General, 130,* 466–478.

Osgood, C. (1952). The nature and measurement of meaning. *Psychological Bulletin, 49,* 172–237.

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 81,* 181–192.

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

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

Plant, E. A., & Devine, P. G. (2000). [Expectations for interracial interactions]. Unpublished raw data.

Plant, E. A., & Devine, P. G. (2001). Responses to other-imposed pro-Black pressure: Acceptance or backlash? *Journal of Experimental Social Psychology, 37,* 486–501.

Plant, E. A., & Devine, P. G. (2002). *Regulatory concerns in interracial interactions: Approaching egalitarianism and avoiding overt bias.* Unpublished manuscript, University of Wisconsin—Madison.

Plant, E. A., Devine, P. G., & Brazy, P. C. (2003). The bogus pipeline and motivations to respond without prejudice: Revisiting the fading and faking of racial prejudice. *Processes and Intergroup Relations, 6,* 187–200.

Rankin, R. E., & Campbell, D. T. (1955). Galvanic skin response to Negro and White experimenters. *Journal of Abnormal and Social Psychology, 51,* 30–33.

Rolls, E. T. (1999). The functions of the orbitofrontal cortex. *Neurocase: Case Studies in Neuropsychology, Neuropsychiatry, and Behavioural Neurology, 5,* 301–312.

Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex, 10,* 284–294.

Rosenthal, R., Rosnow, R. L, & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research: A correlational approach.* New York: Cambridge University Press.

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.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178.

Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology, 57,* 749–761.

Ryan, R. M., Rigby, S., & King, K. (1993). Two types of religious internalization and their relations to religious orientations and mental health. *Journal of Personality and Social Psychology, 65,* 586–596.

Schwartz, G. E., & Kline, J. P. (1995). Repression, emotional disclosure, and health: Theoretical, empirical, and clinical considerations. In J. W. Pennebaker (Ed.), *Emotion, disclosure, and health* (pp. 177–193). Washington, DC: American Psychological Association.

Showers, C. J., Abramson, L. Y., & Hogan, M. E. (1998). The dynamic self: How the content and structure of the self-concept change with mood. *Journal of Personality and Social Psychology, 75,* 478–493.

Smith, E. R. (1993). Social identity and social emotions: Toward new conceptualizations of prejudice. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in group perception* (pp. 297–315). San Diego, CA: Academic Press.

Snyder, M., & Gangestad, S. (1986). On the nature of self-monitoring: Matters of assessment, matters of validity. *Journal of Personality and Social Psychology, 51,* 125–139.

Sutton, S. K., Davidson, R. J., Donzella, B., Irwin, W., & Dottl, D. A. (1997). Manipulating affective stats using extended picture presentations. *Psychophysiology,* 34, 217–226.

van Boxtel, A., Boelhowser, A. J. W., & Bos, A. R. (1998). Optimal EMG signal bandwidth and interelectrode distance for the recording of acoustic, electrocutaneous, and photic blink reflexes. *Psychophysiology, 35,* 690–697.

Vanman, E. J., Boehmelt, A. H., Dawson, M. E., & Schell, A. M. (1996). The varying time courses of attentional and affective modulation of the startle eyeblink reflex. *Psychophysiology, 33,* 691–697.

Vanman, E. J., Paul, B. Y., Ito, T. A., & Miller, N. (1997). The modern face of prejudice and structural features that moderate the effect of cooperation on affect. *Journal of Personality and Social Psychology, 73,* 941–959.

Vidulich, R. N., & Krevanick, F. W. (1966). Racial attitudes and emotional response to visual representations of the Negro. *Journal of Social Psychology, 68,* 85–93.

Vrana, S. R., & Rollock. D. (1998). Physiological response to a minimal social encounter: Effects of gender, ethnicity, and social context. *Psychophysiology, 35,* 462–469.

Vrana, S. R., Spence, E. L., & Lang, P. J. (1988). The startle probe response: A new measure of emotion? *Journal of Abnormal Psychology, 97,* 487–491.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed.). Boston: McGraw-Hill.

Weinberger, D. A., Schwartz, G. E., & Davidson, R. J. (1979). Low-anxious, high-anxious, and repressive coping styles: Psychometric patterns and behavioral and physiological responses to stress. *Journal of Abnormal Psychology, 88,* 369–380.

Whalen, P. J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science, 7,* 177–188.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A., (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience, 18,* 411–418.

Williams, G. C., Rodin, G. C., Ryan, R. M., Grolnick, W. S., & Deci, E. L. (1998). Autonomous regulation and long-term medication adherence in adult outpatients. *Health Psychology, 17,* 269–276.

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

Wundt, W. (1896). *Grundriss der Psychologie* [Outlines of psychology]. Leipzig, Germany: Entgelmann.

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

Zajonc, R. B. (1998). Emotions. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 591–632). Boston: McGraw Hill.

Received November 12, 2001
Revision received June 6, 2002
Accepted June 11, 2002 ■