

## Research Article

# Neural Signals for the Detection of Unintentional Race Bias

David M. Amodio, Eddie Harmon-Jones, Patricia G. Devine, John J. Curtin, Sigan L. Hartley, and Alison E. Covert

University of Wisconsin–Madison

---

**ABSTRACT**—We examined the hypothesis that unintentional race-biased responses may occur despite the activation of neural systems that detect the need for control. Participants completed a sequential priming task that induced race-biased responses on certain trials while electroencephalography was recorded. The error-related negativity (ERN) wave, a component of the event-related potential with an anterior cingulate generator, was assessed to index neural signals detecting the need for control. Responses attributed to race bias produced larger ERNs than responses not attributed to race bias. Although race-biased responses were prevalent across participants, those with larger ERNs to race-biased responses showed higher levels of control throughout the task (e.g., greater accuracy and slowed responding following errors). The results indicate that race-biased responses may be made despite the activation of neural systems designed to detect bias and to recruit controlled processing.

---

Stereotypes of Black people are pervasive in American culture (Gilens, 1996), and as a result, the thoughts and behaviors of White Americans regarding Black people may be influenced by the automatic activation of racial stereotypes (Devine, 1989; Dovidio, Kawakami, & Gaertner, 2002). Although individuals may attempt to prevent racial stereotypes from affecting their behavior by engaging controlled processes, research has demonstrated that intentions to control race bias are not always successful (Devine & Monteith, 1999). Indeed, many self-avowed egalitarians report that prejudices often slip through in their behavior despite their nonprejudiced intentions (Devine, Monteith, Zuwerink, & Elliot, 1991). In some cases, these unintended race biases may lead to pernicious acts of racial discrimination. For example, the stereotype of Black people as violent has been implicated in participants' tendency to more quickly identify a weapon after seeing a Black face than after seeing a White face (Payne, 2001). In subsequent work, participants were faster to "shoot" (via button press) a person carrying an innocuous object (e.g., cell phone) when the person was Black than when the person was White (Correll, Park, Judd, & Wittenbrink, 2002). Such examples of

race-biased discrimination are not limited to laboratory demonstrations. In one highly publicized incident, Amadou Diallo, an African immigrant, was shot dead by New York City police as he reached for an object thought to be a gun: his wallet.

Much research has been focused on elucidating the mechanisms of race-bias control in an effort to understand why control sometimes fails (Devine & Monteith, 1999). A key feature of existing models of prejudice control is that the regulation of behavior follows the conscious realization of having unintentionally made a race-biased response (Monteith, 1993; Monteith, Ashburn-Nardo, Voils, & Czopp, 2002). However, these models do not account for how race-bias activation may be detected and controlled as a response unfolds "in the moment." Recent evidence from the cognitive neuroscience literature has demonstrated that the mechanisms of control are set in motion very early in the response stream and may operate without conscious awareness (Berns, Cohen, & Mintun, 1997; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001). These findings suggest that an understanding of how race-bias control succeeds or fails in ongoing responses requires the examination of rapidly occurring neural processes as an individual is faced with the potential for responding with prejudice. In what follows, we describe a neural model of cognitive control and then propose how it may elucidate the process whereby attempts to control race bias fail.

## NEURAL MODEL OF COGNITIVE CONTROL

Neuroscientists have posited that two separate neural systems work in concert to arrive at an intended behavior in the face of conflicting behavioral tendencies (Botvinick, Braver, Barch, Carter, & Cohen, 2001). The first is a *conflict-detection* system, which monitors ongoing responses and is sensitive to competition between response tendencies. This system is constantly active, requiring few cognitive resources, and has been shown to operate below conscious awareness (Berns et al., 1997; Nieuwenhuis et al., 2001). When conflict is detected by this system, it alerts a second, resource-dependent *regulatory* system designed to implement the intended response while inhibiting the unintended response. This dual-system model offers two explanations for the unintentional expression of race bias. One explanation is that the conflict-detection system does not appraise race-biased response tendencies to be at odds with nonprejudiced beliefs, and therefore the need for control is not signaled. The alternative

---

Address correspondence to David M. Amodio, Department of Psychology, University of California, Los Angeles, 1285 Franz Hall, Los Angeles, CA 90095-1563; e-mail: damodio@ucla.edu.

explanation is that conflict between race-biased behavior and non-prejudiced beliefs is detected and the need for control is signaled, but control is not implemented by the regulatory system.

Support for the dual-system model of control has come from neuroimaging and event-related potential (ERP) studies. Functional magnetic resonance imaging (fMRI) studies have associated the conflict-detection system with activity in the anterior cingulate cortex (ACC), and the regulatory system with activity in the prefrontal cortex (Botvinick, Nystrom, Fissel, Carter, & Cohen, 1999; Carter et al., 1998). Tests of this model have shown that activation of the ACC is high when prepotent response tendencies are at odds with the participant's consciously intended response (MacDonald, Cohen, Stenger, & Carter, 2000). The error-related negativity (ERN) component of the ERP is also sensitive to the conflict-detection process (Gehring & Fencsik, 2001; van Veen & Carter, 2002). The ERN is a negative voltage deflection that typically peaks within 50 ms following a response. It is most strongly pronounced at frontocentral midline scalp sites and has been demonstrated to originate from neural activity located in the ACC (Dehaene, Posner, & Tucker, 1994). ERNs are specifically sensitive to conflicts resulting in errors (i.e., failures to implement control), and larger ERNs have been found to predict subsequent behavioral control, including slowed response latencies and increased accuracy (Gehring, Goss, Coles, Meyer, & Donchin, 1993). ERN indices of conflict detection possess extremely high temporal resolution compared with fMRI, making them appropriate for the examination of conflict detection in rapidly executed responses.

### STUDY OVERVIEW

We examined the role of conflict detection in unintentional expressions of race bias and tested the hypothesis that the neural system for conflict detection is activated when a race-biased response is imminent. To elicit unintentional race-biased responses from participants, we used a sequential priming task adapted from Payne (2001). In this task, participants categorized pictures of either handguns or hand tools following the presentation of White or Black faces. Payne (2001) demonstrated that Black faces facilitated the categorization of "guns," suggesting an automatic association between Black people and weapons. Moreover, when participants were forced to respond quickly (under 500 ms), they were more likely to miscategorize tools as "guns" following Black faces than following White faces. These findings suggest that the correct categorization of "tool" in the context of a Black face requires control over the automatic tendency to associate Black people with weapons. Hence, if the conflict-detection process is particularly sensitive to race-biased response tendencies, it should be activated more strongly on Black-tool trials than on White-tool trials. Consequently, errors on Black-tool trials should produce larger ERNs than errors on White-tool trials.

Payne's (2001) task provides an established method for eliciting race-biased responses and an opportunity to examine the role of conflict detection in race-bias control using ERP methodology. Additionally, it yields multiple indices of automatic and controlled processes. Automatic race bias may be indexed by the degree of response facilitation for stereotype-consistent prime-target pairings (e.g., Black-gun trials), relative to other trial types. Furthermore, if Black faces facilitate the identification of guns, a higher rate of erroneous "gun" responses would be expected when a tool is presented after a Black face compared with a White face. As in past work

(Rabbitt, 1966), control was operationalized as the degree of response slowing and accuracy following errors on the weapons identification task.

Additional indices of automatic and controlled processes in responses on the weapons identification task may be estimated using process dissociation procedures (PDP). The PDP framework assumes that varying levels of automatic and controlled processes contribute to any given response (Jacoby, 1991). The respective contributions of automatic and controlled processes can be dissociated using tasks that place these processes in opposition to one another. For example, when a correct response is congruent with automatic tendencies (e.g., choosing "gun" following a Black face), automatic and controlled processes act in concert. When a correct response is incongruent with automatic tendencies (e.g., choosing "tool" following a Black face), automatic and controlled processes act in opposition. Separate estimates of automatic and controlled processes may be obtained by assessing patterns of hits and misses in participants' performance across congruent (Black-gun) and incongruent (Black-tool) trial types (see Payne, 2001, for PDP formulas).

### METHOD

#### Participants

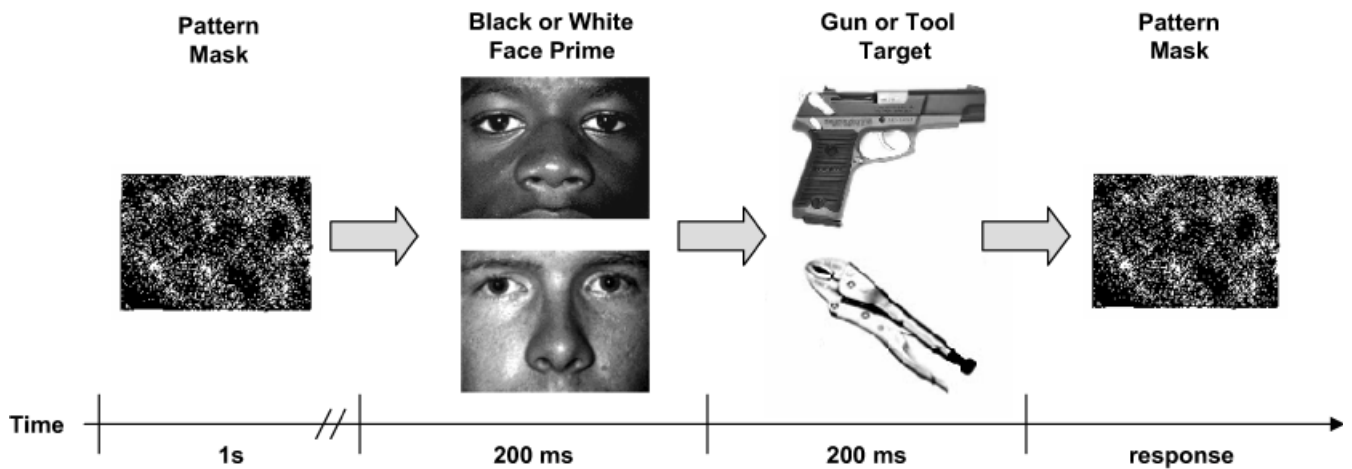
Forty-four right-handed White female students enrolled in an introductory psychology course participated voluntarily for extra course credit. This subject population is notable for holding predominantly nonprejudiced attitudes (Amodio & Devine, 2002). A female sample was used to reduce possible sex-related 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).

#### Procedure

Participants provided informed consent and were then fitted with an electrode cap for electroencephalographic (EEG) recording. Next, the experimenter gave instructions for completing the weapons identification task, explaining that a Black or White face would appear briefly on the computer monitor, followed by a picture of either a handgun or a hand tool. The participants were to classify the second picture as a gun or a tool by pressing a corresponding labeled key with their right or left index finger. Responses were to be made within 500 ms of the target's presentation; participants received a warning to respond more quickly following responses that exceeded this deadline.

The weapons identification task was designed to elicit race-biased responses while minimizing social demand. Participants were told that an erroneous "gun" response on a Black-tool trial was indicative of racial prejudice because it represented an inappropriate application of Black stereotypes. However, participants were assured that their responses would be private and held confidentially, and that their performance could not be monitored by the experimenter. This procedure increased the likelihood that ERNs associated with errors on Black-tool trials could be attributed to intrapersonal concerns about being prejudiced and not to social demand.

At the conclusion of each session, the experimenter probed the participant for suspicion and then explained the hypotheses.



**Fig. 1.** Schematic of the weapons identification task. On a given trial, the prime consisted of either a Black or a White face, and the target consisted of either a handgun or a hand tool (e.g., wrench, drill).

#### Weapons Identification Task

As in the study by Payne (2001), stimuli used in the weapons identification task consisted of two Black and two White male faces, four handguns, and four hand tools (drill, ratchet, wrench, pliers), digitized at  $228 \times 172$  pixels. Stimuli were presented sequentially in the center of a computer screen. Each trial began with a pattern mask (1 s), followed by the prime (200 ms), the target (200 ms), and then a second pattern mask (Fig. 1). The second mask remained on screen until the participant responded or until 2 s had elapsed since the onset of the target. The task consisted of 26 practice trials and 288 experimental trials. Participants completed the weapons identification task while seated approximately 4 feet from a 19-in. monitor refreshing at 100 Hz. Stimuli and recording triggers were presented using DMDX software (Forster & Forster, 2003).

#### EEG Recording and Processing

EEG was recorded from 28 tin electrodes embedded in a stretch-lycra cap (ElectroCap, Eaton, OH), with a left-earlobe reference and forehead ground. Frequencies from DC to 100 Hz were digitized at 2500 Hz using Synamps acquisition hardware (Neuroscan Labs, Sterling, VA). Off-line, EEG was rereferenced to average earlobes, scored for movement artifact, and submitted to a regression-based blink-correction procedure (Semlitsch, Anderer, Schuster, & Presslich, 1986). Frequencies below 1 Hz and above 15 Hz were digitally filtered (96 dB, zero-phase shift). An 800-ms epoch of EEG signal, centered on key press, was selected for each artifact-free trial. For each epoch, baseline-correction procedures subtracted the average voltage occurring 400 to 50 ms before the key press from the entire epoch. ERPs for correct and incorrect trials were averaged within their respective trial types. ERNs were scored as the peak negative deflection occurring between 50 ms before key press and 150 ms after key press at the frontocentral site (Fcz).

## RESULTS

In each analysis, only participants with valid responses on all measures were included. To ensure reliable ERN indices, it was necessary

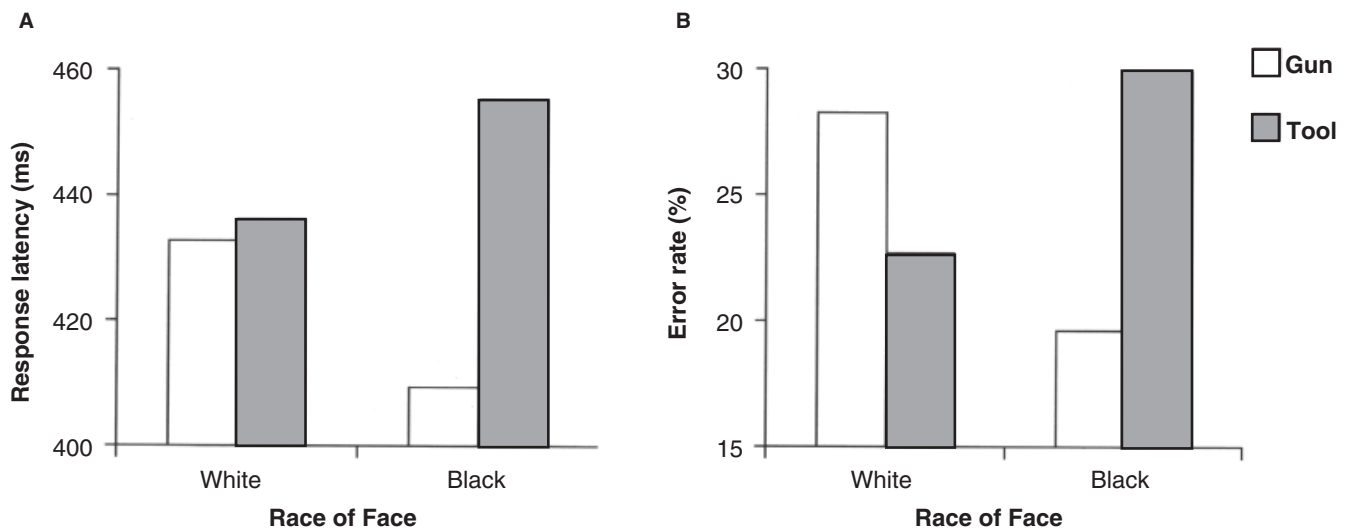
to exclude participants with fewer than five valid ERNs for a given trial type.<sup>1</sup> Thus, the primary analyses were conducted on data from 34 participants.

#### Automatic Race Bias

Behavioral indices of automatic stereotype associations were examined to establish that the task created race-biased response competition. Analyses of reaction times were conducted on correct responses with latencies between 200 and 1,500 ms. Response latencies were log-transformed for analysis, but are reported here in raw milliseconds. A 2 (race of face: White vs. Black)  $\times$  2 (object: gun vs. tool) within-subjects analysis of variance (ANOVA) produced a main effect for object,  $F(1, 33) = 34.04, p < .001$ ; responses to guns ( $M = 420, SD = 36.73$ ) were faster than responses to tools ( $M = 445, SD = 36.13$ ). There was no effect for race of face,  $F < 1$ . The effect for object was qualified by a significant interaction,  $F(1, 33) = 14.39, p < .001$  (Fig. 2a). Simple effect analyses revealed that participants were quicker to identify guns following Black faces than following White faces,  $t(33) = 3.25, p < .005$ , and were slower to identify tools following Black faces compared with White faces,  $t(33) = 2.95, p = .01$ . These results demonstrated an automatic association between Black faces and guns. Additionally, the slower responses to Black-tool pairings suggested that participants adopted a more controlled response strategy on these trials in an effort to avoid errors indicative of race bias.

To determine whether stereotypic associations led to a greater percentage of race-biased errors than nonstereotypic associations did, we conducted a 2 (race of face)  $\times$  2 (object) within-subjects ANOVA for error rates. This analysis produced only a significant interaction,  $F(1, 33) = 10.84, p < .005$  (Fig. 2b). Simple effect analyses revealed that participants made a higher percentage of errors on tool trials when the target was preceded by a Black face, compared with a White

<sup>1</sup>Exclusion resulted either because fewer than five errors were made on a given trial type (1 participant) or because movement artifact rendered EEG signal usable on fewer than five error trials (9 participants on tool trials, 11 on gun trials).



**Fig. 2.** Reaction time (a) and error rate (b) by trial type. Response latencies are presented in raw milliseconds. Shorter latencies indicate a stronger automatic association between the prime (Black vs. White face) and target (gun vs. tool).

face,  $t(33) = 2.92$ ,  $p < .01$ , and on gun trials when the target was preceded by a White face, compared with a Black face,  $t(33) = 2.97$ ,  $p < .01$ . These results indicated a prepotent tendency to favor a “gun” response after viewing a Black face.<sup>2</sup>

#### Neural Signals for the Detection of Race Bias

Analyses of ERNs were conducted separately for tool and gun trials to maximize the number of participants who met the criterion of having five or more valid ERNs per trial type.<sup>3</sup> On average, ERNs consisted of the following number of error-related epochs: 13.91 (Black-tool), 17.97 (White-tool), 13.00 (Black-gun), and 17.03 (White-gun).

ERN amplitudes for tool trials ( $n = 34$ ) were submitted to a 2 (race of face: White vs. Black)  $\times$  2 (response: correct vs. error) ANOVA. A main effect emerged for response,  $F(1, 33) = 114.59$ ,  $p < .001$ ; errors resulted in large negative deflections ( $M = -9.64$ ,  $SD = 5.44$ ), typical of ERNs, whereas correct responses did not ( $M = 0.37$ ,  $SD = 3.03$ ). A main effect for race indicated larger ERN amplitudes for responses to Black-face trials ( $M = -5.05$ ,  $SD = 4.23$ ) compared with White-face trials ( $M = -4.23$ ,  $SD = 4.04$ ),  $F(1, 33) = 4.36$ ,  $p < .05$ . These effects were qualified by a significant interaction,  $F(1, 33) = 10.05$ ,  $p < .005$  (Fig. 3a). Simple effects analysis revealed that ERNs were significantly larger when tools were erroneously classified as “guns” following Black faces ( $M = -10.64$ ,  $SD = 5.74$ ) than when tools were erroneously classified as “guns” following White faces ( $M = -8.64$ ,  $SD = 5.14$ ),  $t(33) = 2.94$ ,  $p < .01$ . ERNs associated with correct “tool” responses following Black and White faces did not differ statistically,  $t(33) = 0.97$ ,  $p = .33$ . The larger ERN observed for errors on Black-tool trials suggests that greater conflict is detected when a response tendency is race biased than when it is not

biased. This finding supports the hypothesis that ACC-related neural processes are sensitive to the potential for a race-biased response, prior to its commission.<sup>4</sup>

The 2 (race of face)  $\times$  2 (response) ANOVA for gun trials ( $n = 32$ ) produced a main effect for response,  $F(1, 31) = 117.80$ ,  $p < .001$ ; erroneous classifications of guns as “tools” produced large ERNs ( $M = -8.84$ ,  $SD = 5.07$ ), whereas correct responses did not ( $M = 0.21$ ,  $SD = 2.84$ ). Race of face was not significant,  $F < 1$ . The interaction was not significant,  $F(1, 32) = 1.12$ ,  $p = .30$  (Fig. 3b), indicating that the amplitude of ERNs related to erroneous “tool” responses did not differ statistically when targets followed Black faces ( $M = -8.66$ ,  $SD = 4.74$ ) compared with White faces ( $M = -9.01$ ,  $SD = 5.40$ ). As in the study by Payne (2001), Black-gun trials were not expected to elicit heightened conflict because “gun” responses were required by the task and would not have been attributable to race bias.

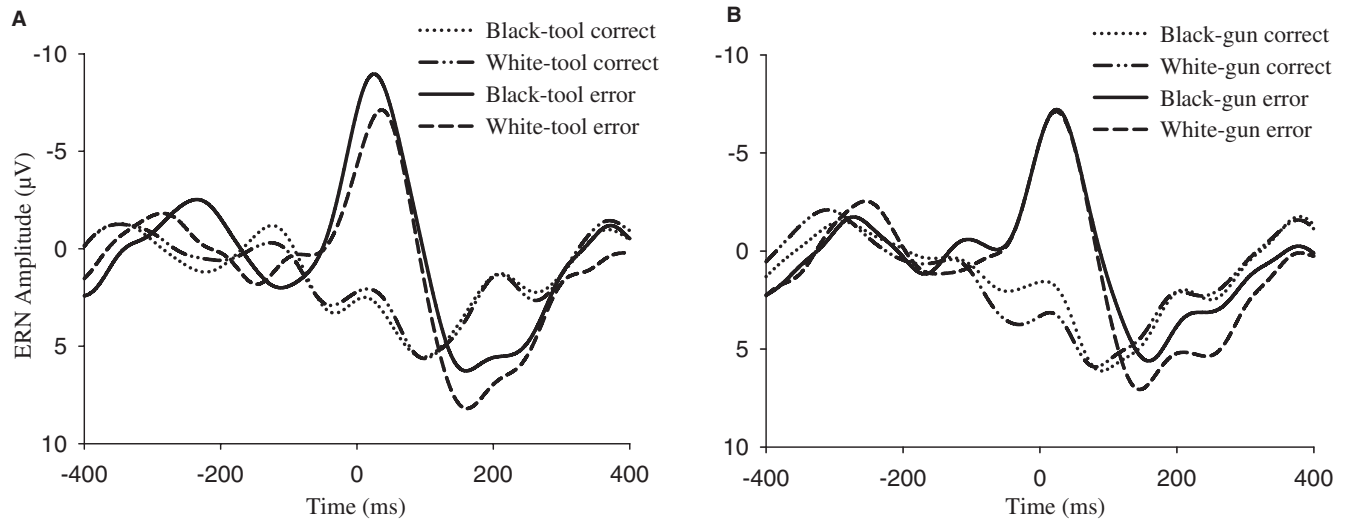
#### Neural Signals of Race-Bias Detection Predict Response Control

Correlational analyses explored the relationship between participants’ ERNs and indices of behavioral control. To examine the unique effects of ERNs associated with unintentional race bias, we computed

<sup>2</sup>When all 44 participants were included, similar Race of Face  $\times$  Object interactions were observed for response latency,  $F(1, 43) = 17.82$ ,  $p < .001$ , and error rate,  $F(1, 43) = 11.43$ ,  $p < .005$ .

<sup>3</sup>Although only 29 participants met the inclusion criterion for the full 2 (race of face)  $\times$  2 (object)  $\times$  2 (response) ANOVA, this analysis produced the predicted three-way interaction,  $F(1, 28) = 4.24$ ,  $p < .05$ .

<sup>4</sup>An ERN-like deflection occurring on correct responses has been referred to as the N2 component. Like the ERN, the N2 is associated with ACC activity, but it reflects conflict-detection processes leading to successful control (van Veen & Carter, 2002). In this study, a larger N2, indicating high conflict, would be expected for correct responses on Black-tool trials compared with correct responses on other trial types. N2 amplitudes were quantified as the peak negative voltage deflection at Fcz between 250 ms and 450 ms after target (gun or tool) onset, on correct trials. A planned comparison revealed that N2 amplitudes for Black-tool trials ( $M = -5.70$ ,  $SD = 2.61$ ) were significantly larger than N2s averaged among the remaining three trial types ( $M = -4.17$ ,  $SD = 2.25$ ),  $t(33) = -5.69$ ,  $p < .001$ . In pair-wise comparisons, Black-tool N2s were larger than N2s for each other trial type,  $t_s > 3.50$ ,  $p_s < .001$ . This pattern remained significant when data from all 44 participants were included. It should be noted, however, that because the N2 and ERN are quantified differently (e.g., via stimulus-locked vs. response-locked averaging), their amplitudes are not directly comparable.



**Fig. 3.** Response-locked event-related potential waveforms for correct and incorrect tool (a) and gun (b) trials as a function of race of face. Zero indicates the time of the key press. Errors on Black-tool trials represent race-biased responses. ERN = error-related negativity.

a *race-bias ERN* as the residual variance of the Black-tool ERN amplitude after White-tool ERN amplitude was covaried. Thus, the race-bias ERN reflected variance due to race (Black vs. White faces) and not to the target (tool).

Larger (negative-polarity) race-bias ERNs were associated with longer response latencies on trials following errors,  $r(32) = -.44$ ,  $p = .01$ . The correlation between race-bias ERNs and posterror response slowing was significant for each trial type,  $r_s > -.33$ ,  $p_s < .055$ , suggesting that participants with larger race-bias ERNs responded more carefully on all trial types than participants with smaller race-bias ERNs. Larger race-bias ERNs were also associated with greater accuracy following errors,  $r(32) = -.50$ ,  $p < .005$ . An analysis by target type, averaging across race of face, revealed that race-bias ERNs were significantly correlated with posterror accuracy on tool trials, which carried the potential for race bias,  $r(32) = -.48$ ,  $p = .005$ , but not on gun trials, which did not carry the potential for race bias,  $r(32) = -.27$ ,  $p = .13$ . Although the difference between these correlations was marginal,  $t(31) = 1.43$ ,  $p = .08$  (one-tailed), the correlation between race-bias ERNs and posterror accuracy on tool trials remained significant when posterror accuracy on gun trials was partialled out,  $r(31) = -.44$ ,  $p < .01$ .<sup>5</sup> These results demonstrate that greater activity of ACC-related systems in response to race-biased response conflict was associated with a fairly specific pattern of control.

#### Process Dissociation Analysis

Using PDP formulas, we calculated estimates of automatic and controlled processes separately for White- and Black-face trials, as in Payne (2001). A 2 (race of face: White vs. Black)  $\times$  2 (process:

automatic vs. controlled) within-subjects ANOVA produced a significant interaction,  $F(1, 33) = 9.33$ ,  $p < .005$ . Simple effect analysis revealed that automatic processes contributed more to response following Black faces ( $M = .59$ ,  $SD = .15$ ) than to response following White faces ( $M = .47$ ,  $SD = .15$ ),  $t(33) = 3.23$ ,  $p < .005$ . Estimates of controlled processes did not differ by race of face,  $t(33) = 0.38$ ,  $p = .71$ .<sup>6</sup>

Larger race-bias ERNs were associated with higher PDP estimates of control on trials involving both Black faces,  $r(32) = -.46$ ,  $p < .01$ , and White faces,  $r(32) = -.40$ ,  $p < .03$ , consistent with the previously reported relationship between race-bias ERNs and the other indices of control. Race-bias ERNs were not correlated with PDP estimates of automatic processes in responses to Black faces,  $r(32) = .19$ ,  $p = .28$ , or to White faces,  $r(32) = .24$ ,  $p = .17$ .

#### DISCUSSION

Our results demonstrate that neural mechanisms of conflict detection are sensitive to response competition involving automatically activated racial stereotypes. Given that race-biased response tendencies are met with enhanced signals for control, one may wonder how unintentional race biases make their way into people's everyday behavior. Typically, unintended race bias occurs when responses are made quickly and in the absence of sufficient processing resources (Amodio, Harmon-Jones, & Devine, 2003; Gilbert & Hixon, 1991). This explanation is consistent with our finding that participants took more time to respond correctly to Black-tool pairings than to any other pairing,  $t_s > 2.95$ ,  $p_s < .01$ . That is, participants were more willing to risk exceeding the response deadline when an incorrect response

<sup>5</sup>A similar pattern of correlations was obtained for error rates when Black-gun error ERNs were covaried from Black-tool error ERNs; this alternative race-bias ERN was associated with fewer tool errors,  $r(31) = .54$ ,  $p = .002$ , but uncorrelated with gun errors,  $r(31) = .14$ ,  $p = .47$ .

<sup>6</sup>The Race of Face  $\times$  Process interaction for PDP estimates was replicated when data from the full sample were used,  $F(1, 43) = 6.90$ ,  $p < .02$ . Simple effects revealed a higher estimate of automatic processing in responses to Blacks compared with responses to Whites,  $t(43) = 3.10$ ,  $p < .005$ . Estimates of controlled processes did not differ by race of face,  $t(43) = 0.89$ ,  $p = .38$ .

would have indicated racial prejudice than when it would not. Taken together with our other results, this finding suggests that despite an enhanced signal for control to prevent race bias, the successful implementation of control may be impeded by lack of time or cognitive resources (e.g., divided attention).

The demonstration that conflict-detection processes are activated in the face of race-biased tendencies at very early stages of response execution suggests an extension of extant models of prejudice control (e.g., Monteith, 1993) and mental correction (Wegener & Petty, 1997; Wilson & Brekke, 1994). These models propose that race-bias detection is associated with the conscious awareness that race bias has been committed. Although our results support the key role of race-bias detection in controlling unwanted responses, they suggest that the detection process may operate below awareness and therefore does not necessarily rely on conscious reflection. Further research is needed to more fully explore the role of conflict detection in race-bias control and to examine the extent to which individual differences in the detection process may predict the ability to regulate expressions of prejudice.

The present investigation of race-bias conflict detection extends previous neurophysiological research by addressing responses that are personally and socially relevant, and for which errors represent significant personal transgressions. This line of research suggests an opportunity for future investigations of how neural mechanisms of automaticity and control may be moderated by situational and personal factors, highlighting the utility of exploring the functions of the brain in its social context.

**Acknowledgments**—We thank Keith Payne for providing stimuli for the weapons identification task. This research was supported by National Institute of Mental Health Training Grant T32-MHI8931, a University of Wisconsin dissertation fellowship, and an American Psychological Association dissertation research award to D.M.A.; a Wisconsin Alumni Foundation grant to P.G.D.; and a National Science Foundation grant (BCS-9910702) to E.H.-J.

## REFERENCES

- Amodio, D.M., & Devine, P.G. (2002). [General attitudes questionnaire, 1999–2001]. Unpublished raw data.
- Amodio, D.M., Harmon-Jones, E., & Devine, P.G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink responses and self-report. *Journal of Personality and Social Psychology, 84*, 738–753.
- Berns, G.S., Cohen, J.D., & Mintun, M.A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science, 276*, 1272–1275.
- Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., & Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652.
- Botvinick, M.M., Nystrom, L.E., Fissel, K., Carter, C.S., & Cohen, J.D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402*, 179–181.
- Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., & Cohen, J.D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*, 747–749.
- Correll, J., Park, B., Judd, C.M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83*, 1314–1329.
- 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.
- Dehaene, S., Posner, M.I., & Tucker, D.M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science, 5*, 303–305.
- Devine, P.G. (1989). Prejudice and stereotypes: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5–18.
- 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., & Monteith, M.M. (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.
- Dovidio, J.F., Kawakami, K., & Gaertner, S.L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology, 82*, 62–68.
- Forster, K.I., & Forster, J.C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116–124.
- Gehring, W.J., & Fencsik, D.E. (2001). Functions of the medial frontal cortex in the processing of conflict and errors. *Journal of Neuroscience, 21*, 9430–9437.
- Gehring, W.J., Goss, B., Coles, M.G.H., Meyer, D.E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science, 4*, 385–390.
- Gilbert, D., & Hixon, J.G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*, 509–517.
- Gilens, M. (1996). Race and poverty in America: Public misperceptions and the American news media. *Public Opinion Quarterly, 60*, 515–541.
- Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.
- MacDonald, W., III, Cohen, J.D., Stenger, V.A., & Carter, C.S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288*, 1835–1838.
- Monteith, M.M. (1993). Self-regulation of stereotypical responses: Implications for progress in prejudice reduction. *Journal of Personality and Social Psychology, 65*, 469–485.
- Monteith, M.M., 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.
- Nieuwenhuis, S., Ridderinkhof, K.R., Blom, J., Band, G.P.H., & Kok, A. (2001). Error-related brain potentials are differently related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology, 38*, 752–760.
- 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.
- Rabbitt, P.M.A. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology, 71*, 264–272.
- Semlitsch, H.V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology, 23*, 695–703.
- van Veen, V., & Carter, C.S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience, 14*, 593–602.
- Wegener, D.T., & Petty, R.E. (1997). The flexible correction model: The role of naïve theories of bias in bias correction. In M.P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 141–208). Mahwah, NJ: Erlbaum.
- Wilson, T.D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116*, 117–142.

(RECEIVED 9/23/02; REVISION ACCEPTED 2/27/03)