

Standing Up for a Change: Reducing Bias Through Interpersonal Confrontation

Alexander M. Czopp
University of Toledo

Margo J. Monteith and Aimee Y. Mark
University of Kentucky

Three experiments examined the effectiveness of interpersonal confrontations as a means for decreasing stereotypic responding. After making stereotypic inferences about Black individuals, participants were confronted and reactions were measured across various intrapersonal and interpersonal response domains. Confrontations varied in level of hostility (Experiment 1) and whether they were expressed by a Black or White person (Experiment 2). Results indicate that although confrontations (and particularly hostile ones) elicited negative emotions and evaluations toward the confronter, participants also experienced negative self-directed affect. Furthermore, regardless of who did the confronting or how much hostility was expressed, confronted participants subsequently were less likely to provide stereotypic responses (Experiments 1–2), and the effect of the confrontation generalized to reporting less prejudiced attitudes (Experiment 3).

Keywords: interpersonal confrontation, prejudice reduction, racism, stereotypes

Bob Dylan's song *Oxford Town* (Dylan, 1963/2004) describes the struggle of James Meredith, the first African American to be admitted to the University of Mississippi in 1962. Meredith's admittance sparked a fire of controversy, conflict, and chaos. Nonetheless, his admittance spurred the desegregation of the school, and Meredith's successful graduation in 1964 provides a powerful example of how one person can confront and overcome personal, social, and institutional injustices. Indeed, much of the sociopolitical activism of the Civil Rights Movement (e.g., Rosa Parks's refusal to give up her seat to a White man on a Montgomery, AL, bus) can be characterized as confrontations against institutional prejudice and discrimination.

Opportunities for confrontation occur not only on such a grand scale but frequently take place in everyday life. For example, imagine a situation where a group of friends or family is gathered together and during the conversation someone in the group says something that raises a few eyebrows. Maybe it is a comment

about how "Blacks need to get off welfare" or perhaps something more subtle like a racial joke intended to elicit a chuckle from the group. What is a bystander who hears such comments and finds them offensive to do in such a situation? Some people would rather remain silent and not want to create a stir, but perhaps creating a stir can have an impact and induce change in others. By speaking up and addressing the person who made prejudiced comments and voicing displeasure, people have the opportunity to confront prejudice directly and perhaps help reduce the likelihood of such aspersions in the future.

Confronting others may be a strategy for directly or indirectly influencing a variety of intergroup biased responding. For example, confrontations may induce changes in people's endorsement of negative stereotypes of outgroups or their underlying affective–evaluative reactions toward outgroup members (i.e., prejudice). Even if confrontations merely inhibit another person's immediate and public expressions of prejudice rather than any underlying beliefs, they may have beneficial consequences by creating norms of egalitarianism and instigating self-regulatory processes. However, within the stereotyping and prejudice literature, empirical research examining the efficacy of interpersonal confrontation as a prejudice-reduction strategy is surprisingly sparse. The goal of the present research was to conduct several experiments examining people's reactions when they are confronted by another person to determine the potential role of interpersonal confrontations in curbing a variety of group-based, biased responses.

Bias Reduction Through Confrontation

Racial bias is a specific form of category-based responding that includes (typically) negative affect, stereotypes, attitudes, or behaviors directed toward members of racial outgroups (e.g., African Americans). The notion that confrontations can effectively induce changes in racially biased responding was suggested in the 1970s. Rokeach (1973; Rokeach & Cochrane, 1972; Rokeach & McLellan, 1972) demonstrated that when individuals were confronted

Alexander M. Czopp, Department of Psychology, University of Toledo; Margo J. Monteith and Aimee Y. Mark, Department of Psychology, University of Kentucky.

Experiment 1 of this article is based on a thesis submitted by Alexander M. Czopp to the University of Kentucky in partial fulfillment of the requirements for the master's degree.

Portions of this research were presented at the annual meetings of the Society for Personality and Social Psychology, Savannah, Georgia, February 2002, and Los Angeles, California, February 2003. We thank committee members Monica Harris and Richard Smith for their constructive advice regarding Experiment 1; Leslie Ashburn-Nardo for helpful comments on earlier drafts of this article; and Michael Dudley, Jennifer Gardner, Leslie Johnson, and Chloe Steinbrunner for their assistance in data collection.

Correspondence concerning this article should be addressed to Alexander M. Czopp, Department of Psychology, University of Toledo, 2801 West Bancroft Street, Toledo, OH 43606. E-mail: alexander.czopp@utoledo.edu

with the fact that their egalitarian self-concept was inconsistent with their prejudiced values, attitudes, and behaviors, they experienced feelings of self-dissatisfaction. To reduce the negative affect associated with these contradictions, individuals often changed their attitudes and behaviors to be more consistent with their self-concepts. Rokeach's research using the value self-confrontation method suggests that the technique can induce long-term change (i.e., up to 21 months after self-confrontation) in attitudes (e.g., more favorable attitudes toward civil rights) and behaviors (e.g., joining the National Association for the Advancement of Colored People). More recently, researchers have found that awareness of discrepancies between one's behavior (i.e., having prejudicial tendencies) and one's self-concept (i.e., being a fair and egalitarian person) gives rise to feelings of negative, self-directed affect (Devine, Monteith, Zuwerink, & Elliot, 1991; Monteith & Voils, 1998). These feelings of self-dissatisfaction and guilt prompt the activation of a self-regulatory cycle that encourages more egalitarian and self-consistent responding in subsequent situations (Monteith, 1993; Monteith, Ashburn-Nardo, Voils, & Czopp, 2002).

Thus, self-induced confrontations may be an effective means of reducing one's own racially biased responding. However, for whom will self-induced confrontations be most effective and under what circumstances? Research suggests that the success of self-confrontation approaches relies heavily on both awareness and motivation. That is, people must be able to recognize their biased responses, and they must be motivated to correct them. Increasing evidence suggests that, because of the pervasiveness of stereotypes and the automatic nature of stereotype activation, people will not always recognize their tendencies for biased responding (Bargh, 1999; Devine & Monteith, 1999). In addition, Monteith's (1993) research suggests that, after becoming aware of a prejudiced response, only individuals with low-prejudiced personal standards experience self-dissatisfaction and subsequently try to control their prejudiced responses. In contrast, more prejudiced individuals find their prejudiced responses acceptable and appropriate and are consequently less motivated to change or control such responses. In summary, to the extent that people are either unaware of their own prejudiced responses or are unmotivated to change such responses, self-confrontation strategies may not operate successfully.

This is where confrontations from others may play a role. Despite little research examining other-induced confrontations, there is reason to believe that such interpersonal exchanges may be effective in reducing biased responding. Among low-prejudice people, interpersonal confrontations may create awareness of one's biased responding, thereby eliciting feelings of self-dissatisfaction and prompting subsequent self-regulation (Monteith, 1993). In addition, such confrontations may be a means of creating situational norms that underscore principles of fairness and egalitarianism that both low- and high-prejudice individuals have been socialized to embrace (Gaertner & Dovidio, 1986; Monteith & Walters, 1998). Previous research suggests that when norms of fairness and egalitarianism are made salient, people (regardless of their prejudice level) become less likely to provide prejudiced responses (Blanchard, Lilly, & Vaughn, 1991; Monteith, Deneen, & Tooman, 1996). However, because interpersonal confrontations can be awkward or controversial, are people actually likely to confront others' prejudice?

Work by Alicke et al. (1992) that examined the interpersonal functions of complaining behavior suggested that people frequently use complaints as attempts to change the attitudes or behaviors of others. Within the realm of prejudice, Feagin (1991) recorded Black individuals' various responses to public experiences of racial discrimination and found that the most frequent response to the incident was to confront the source of discrimination (see also Swim, Hyers, Cohen, Fitzgerald, & Bylsma, 2003). Similarly, Swim and Hyers (1999) examined the responses of female participants to a male confederate who made several sexist comments and found that almost half the women confronted the confederate and that the women most likely to confront were those firmly committed to fighting sexism. These examples suggest that confrontation is indeed an interpersonal strategy used by people in their efforts to fight prejudice. Below, we consider factors that are likely to influence how interpersonal confrontations are likely to be received.

Hot Versus Cold Confrontations

When one thinks of interpersonal confrontations, images of furrowed brows and clenched teeth come to mind. Indeed, confrontations involving prejudice may erupt into very contentious interactions. How an individual reacts to being confronted about a prejudiced response will likely be determined in part by the manner in which the issue is broached. Hostile and accusatory confrontations not only may violate social norms of politeness and decorum, they may impugn the confronted person's self-concept as a good and decent individual. Low-prejudice individuals' personal standards are closely tied to their self-concepts (Devine et al., 1991). To them, a hostile allegation that they are prejudiced represents a firm blow to their positive and nonprejudiced self-image. Conversely, although high-prejudice people do not have such internalized standards for nonprejudice, they recognize the social stigma associated with being publicly labeled a bigot. In a democratic society that emphasizes egalitarianism, a racist or sexist designation carries serious negative connotations.

Self theorists have underscored the negative consequences of threats to one's self-image. For example, Tangney and her colleagues (Tangney, 1995; Tangney, Wagner, Fletcher, & Gramzow, 1992) pointed out that feelings of shame associated with a negative evaluation from another person can easily turn into anger directed toward that person. On the basis of the findings of Baumeister, Smart, and Boden (1996), Baumeister and Campbell (1999) argued that "when people feel that their favorable self-images or reputations have been impugned by someone, they may become motivated to attack that person in a violent or aggressive fashion" (p. 218). Thus, to the extent that an interpersonal confrontation poses a threat to the confronted individual's favorable self-concept, reactions are likely to be negative.

It is interesting that although hostile confrontations are likely to lead to negative reactions toward the confronter, they may nonetheless be effective in reducing future biased responses. Despite disliking the confronter, people may walk away wary of subsequent situations in which a prejudiced response is possible. Among low-prejudice people, a racially biased response (regardless of how it was acknowledged) violates personal standards and, thus, may lead to guilt and the self-regulation of prejudiced responses. Even more prejudiced individuals, many of whom are externally moti-

vated to avoid being perceived by others as prejudiced (Fazio, Jackson, Dunton, & Williams, 1995; Plant & Devine, 1998), may be less likely to engage in biased responding in the future to avoid such public dispute or disapproval. Thus, a confronter may be able to accomplish the goal of discouraging future biased responding even with a hostile confrontation. However, individuals often seek to establish, maintain, or enhance social relationships. Can confrontations reduce others' biased responses and not have a negative impact on interpersonal relationships? Perhaps less threatening confrontations that emphasize universally accepted and embraced principles of fairness and egalitarianism would be effective in curbing bias and circumventing negative interpersonal consequences.

Black Versus White Confrontations

A second important factor that is likely to influence people's reactions to being confronted about a biased response is the group membership of the confronter. Being confronted by a member of the group that one has just offended (e.g., Blacks, women) is likely to yield different reactions than being confronted by a nontarget member (e.g., Whites, men). Recent research suggests that confrontations from target group members evoke different reactions than similar confrontations from nontarget individuals. Using a scenario-based approach in which participants imagined various confrontations, Czopp and Monteith (2003) found that participants indicated they would feel more guilty and be more likely to apologize when confronted about a prejudiced response by a White person than a Black person (in the case of racism) and by a man than a woman (in the case of sexism). Czopp and Monteith argued that people may expect targets to be particularly likely to confront others on the basis of their self-interest in curbing biased responding directed toward their group. Persuasion research suggests that when individuals argue in favor of their group's interests, others are less likely to process their message; in contrast, individuals who argue against their group's interest often surprise people and their messages receive greater processing (Petty, Fleming, Priester, & Feinstein, 2001). Because a nontarget group member's confrontation violates expectancies of who is most likely to confront prejudice, a nontarget group member may evoke greater message processing and, to the extent that the confrontation includes strong arguments, greater feelings of negative, self-directed affect (e.g., self-criticism) and ultimate persuasion. An alternative possibility is that target confronters are perceived as hypersensitive and as complainers (see Kaiser & Miller, 2001, 2004), and such perceptions ultimately render their confrontations less effective. Recent findings reported by Mark, Monteith, and Czopp (2005) supported this complainer mediational hypothesis rather than the expectancy-violation mediational hypothesis.

Thus, previous research using imagined situations suggests that nontarget group members may be more effective confronters than target group members. However, would similar findings emerge in actual situations where two people are interacting with each other and one is confronted by the other person? Unlike hypothetical scenarios, a real interpersonal confrontation may create a situational dynamic in which self-presentation concerns become paramount. Confronted individuals may be especially concerned about how they are perceived by the confronting individual (e.g., as bigots). Because of the salience of such interpersonal concerns,

people may be particularly sensitive to the potential negative evaluation from a target group member. That is, being perceived as racist by a Black person may be more unsettling than similar perceptions by a White person. Indeed, recent research by Winslow (2004) involving actual dyadic interactions suggested that White persons are more concerned about imputations of prejudice from a Black person than another White person. If this applies to situations involving interpersonal confrontations, people may experience greater negative self-directed affect in response to a target member's confrontation than a nontarget's confrontation.

The importance of examining reactions to confrontations from target group members is underscored further by research suggesting that some targets are reluctant to confront others' biased responding because of possible deleterious interpersonal consequences. Although almost half of the women in Swim and Hyers' (1999) study confronted a male confederate's sexist remarks, many participants failed to confront the confederate out of fear of being perceived negatively. Similarly, Kaiser and Miller (2004) found that many women experience great displeasure during interactions involving sexism but fail to actively confront others out of concern about incurring negative interpersonal costs (see also Shelton & Stewart, 2004). Whereas remaining silent and not confronting may circumvent interpersonal costs, recent findings suggest that failure to confront is associated with negative intrapersonal consequences. Shelton, Richeson, Salvatore, and Hill (2006) found that women who believed they should confront sexist treatment but did not do so experienced negative self-directed affect. However, the studies discussed above examined reactions to prejudice from the confronter's perspective rather than the perpetrator's perspective. Examining the reactions of confronted individuals is essential for understanding the appropriateness of targets' apprehension in confronting prejudice. If such confrontations do not yield negative interpersonal reactions, targets may be more likely to use confrontation as an interpersonal strategy for reducing bias. In addition, even if confrontations result in negative evaluations, to the extent that they are effective in decreasing future biased responses, targets may be willing to sacrifice positive social interactions in the name of curbing bias.

The Current Experiments

In the present research, we focused on how White individuals respond to confrontations of bias related to African Americans. We chose this particular context because prejudice-related responding directed toward African Americans is a sensitive topic, and race-related confrontations are likely to be taken more seriously than those involving other social groups (e.g., women, older persons, gay and lesbian individuals; see Rodin, Price, Bryson, & Sanchez, 1990). As a result, people's reactions to confrontation are likely to be quite variable. We studied participants' reactions to confrontation across three experiments, and in each experiment we assessed a variety of response domains. First, the interpersonal nature of a confrontation requires that the confronted person provide some kind of direct public response, so we assessed the immediate reaction in the situation. Second, on an intrapersonal level, because people may have a variety of emotional reactions to being confronted, ranging from anger to embarrassment, we assessed participants' affective reactions. Third, because someone else is necessarily involved in an interpersonal confrontation, the confronted

individual's perception and evaluation of the confronter were considered. Last, it is important to consider the effect of confrontation on the person's future responses (i.e., will biased responding decrease?).

In the first experiment, we examined White participants' reactions to confrontations of racially biased responding that varied in hostility. In Experiment 2, we investigated the possible influence of the confronter's race across the various response domains. In Experiment 3, we examined the generalization of the effects of confrontation. Specifically, in the first two experiments we examined whether confrontation about stereotypic responses subsequently resulted in less stereotypic responses on a similar type of task. In contrast, in Experiment 3, we investigated whether confrontation also affected prejudice-related responding on tasks that were different from the one used in the confrontation situation.

Experiment 1

In Experiment 1, we compared reactions between a low-threat confrontation that emphasized norms of fairness and a more self-threatening and accusatory confrontation. We expected that a threatening confrontation that makes an accusation of racism would lead to more negative immediate reactions and more negative affect directed at the confronter relative to a less threatening confrontation. In contrast, we expected that the less threatening confrontation that appeals to universally embraced ideals of fairness would yield more positive immediate reactions and more favorable evaluations of the confronter relative to the more threatening confrontation. However, because both types of confrontations likely increase the salience of personal and social norms for responding without prejudice, we expected that both types of confrontations would lead to less stereotypic responding.

Method

Participants

One hundred twenty-one White participants (60 men, 61 women) were recruited from the department's introductory psychology participant pool and received course credit. Potential participants had completed the 20-item Attitudes Toward Blacks scale (ATB; Brigham, 1993) as part of a larger survey at the beginning of the semester. For recruitment purposes, the distribution of ATB scores was dichotomized (on the basis of a median split) and an equal number of students from each side of the distribution were randomly identified as potential participants to ensure a wide range of prejudice scores. (However, ATB scores were treated as a continuous variable in all analyses.) Participants were contacted by telephone and asked to participate but were not informed of the relation between selection criteria and the present research. Both the person contacting potential participants and the experimenter were blind to participants' ATB scores.

Overview

Participants believed they were part of a dyad, working together in separate rooms via networked computers on a project studying people's attitudes toward computers and computer use. To simulate the various ways people use computers for personal and professional use, each participant and his or her "partner" (actually the experimenter posing as the other participant) completed two tasks. The first task resembled an online conversation between two people getting to know each other and was simply a filler task. The second task required participants to make inferences about people on the basis of a photograph and a one-sentence description. Some

sentences yielded inferences that, when made about a Black person, were consistent with common stereotypes of Black individuals. After participants completed three of such potentially biased sentences, they were confronted by the confederate in either a threatening or nonthreatening way. Participants' immediate typewritten responses to the confrontation were recorded. Afterward, participants provided affect and partner evaluation ratings and were given a second opportunity to respond privately to photograph-sentence pairs that were similar to those completed with their partner.

Procedure

Participants completed the experiment individually but were told that they would be working with a partner. On arrival at the designated room, the experimenter informed the participant that because there were two people completing this study, instructions would be delivered through an intercom system from another room so that both participants could be addressed simultaneously. Depending on when the participant arrived for the experiment, the experimenter informed him/her that the "other subject" was running a little late or had already arrived.¹

From a nearby room that was also equipped with an intercom and a networked computer, the experimenter addressed participants about the purpose of this study. Participants completed two tasks: The first task was intended to represent a social interaction similar to how people get to know each other through e-mail or chat rooms, and the second task represented a more instrumental use of computer communication involving two people working together to complete a project. Participants were told that they would be working with the other person exclusively over the computer and they would not meet or interact face to face.

Participants communicated with each other using America Online's (AOL) Instant Messenger program (America Online, 2000), which allows two people to "talk" to each other by typing their responses back and forth. This program also stores the entire dialogue on its screen making it possible for the experimenter to copy and paste the entire transcription of each participant. After familiarizing participants with this program, participants exchanged information about themselves with the confederate. To facilitate this conversation (and allow for participants to receive standardized information from the confederate), the experimenter instructed participants to follow a specific order and include specific topics such as their personal life, hobbies, and academic background. The confederate's information was scripted and held constant across all participants. The confederate gave no indication of gender but provided information to suggest being White (i.e., mentioning a European family background).

The second task was intended to simulate how two people might work together to complete a project by communicating via computer only. The project involved looking at photographs of people, reading sentences, and making inferences about people on the basis of those sentences. This task was similar to one used by Monteith et al. (2002, Experiment 3). The photographs were of Black and White men and women. The sentences were constructed so that an inference could be made about the photographed person on the basis of the sentences. For example, the sentence, "This person works with numbers," could yield responses such as "accountant" or "math teacher." Also included were photograph-sentence descriptions intended to yield negative inferences (e.g., "This person steals money"). These stimuli were presented in a notebook in the participant's room.

Participants alternated with the confederate in providing responses for each photograph-sentence pair, but both the participant and the confederate

¹ Throughout the experiment, the same person was playing the role of the experimenter and the second participant (i.e., the confederate). To clarify which role the experimenter assumed, *confederate* indicates performing the role of the other participant and *experimenter* indicates performing the duties of the experimenter.

provided a response to each stimulus. The experimenter instructed participants that they were to provide responses that made the most sense to them, but their responses must be different from the response given by their partner. As in the conversation section, the confederate provided scripted responses.

Included among 13 filler items were three critical photograph–sentence pairs. The sentences for these items yielded responses that, because they were paired with a photograph of an African American person, could be construed as stereotypic in nature. For example, the sentence, “This person depends on the government for money,” is likely to yield responses such as “poor person” or “on welfare.” When such an inference is made about a Black person, the response is consistent with stereotypes about Blacks and could be perceived as racially biased. The other two critical sentences were, “This person can be found behind bars,” and, “This person can be found wandering the streets,” and could yield stereotypic inferences such as “criminal” and “homeless person,” respectively. The critical items were strategically located in the task such that the participant was always the first to respond to them. The confederate always provided nonstereotypic alternative responses (e.g., “federal employee” instead of “welfare recipient,” “bartender” instead of “criminal,” “tourist” instead of “bum”).

After participants responded to the third critical photograph–sentence pair (which was also the last item for the task), the confederate quickly typed a nonstereotypic response and one of two messages that served as the confrontation. Both confrontations began by pointing out the participant’s responses: “By the way, for some of the last pictures of Black people, you said things like bum, person on welfare, and criminal. i know these things make sense based on the descriptions we were given.” The actual wording of the confrontation was tailored for each participant so as to include the exact stereotypic response(s) provided by the participant. The confederate then continued by typing one of the following confrontation messages:

Low threat: but maybe it would be good to think about Blacks in other ways that are a little more fair? it just seems that a lot of times Blacks don’t get equal treatment in our society. you know what i mean?

High threat: but you should really try to think about Blacks in other ways that are less prejudiced. it just seems that you sound like some kind of racist to me. you know what i mean?

A pilot study was conducted to test the threat level of these confrontation messages. Forty-four participants imagined they had made stereotypic inferences to the photograph–sentence task described above and were subsequently confronted in either a low- or high-threat manner, using the exact wording used in our two threat conditions. Participants then indicated the degree to which they felt the confrontation was a nonthreatening request (e.g., “I feel the confrontation was a casual, nonthreatening suggestion”) or a hostile personal attack (e.g., “In this confrontation, my partner is calling me a bad person”). The six questions displayed high internal consistency ($\alpha = .89$). A composite score of the six items was created, with lower scores indicating a less threatening confrontation and higher scores representing a more threatening confrontation. As expected, the low-threat confrontation ($M = 2.71$) was rated significantly lower than the high-threat condition ($M = 4.17$), $t(42) = 3.48$, $p = .001$, $r = .47$.

After typing the confrontation, the confederate allowed sufficient time to pass for a written reply by the participant. After the participant’s response, the confederate pressed a button on the intercom to indicate to the experimenter that this task was finished, and the experimenter informed participants that the interactive part of the experiment had ended. At this point, the experimenter entered the participant’s room and explained that they would no longer be interacting with the other student in any way. In addition, the experimenter explained that all responses the participant made from this point on would be recorded confidentially and that neither the other student nor the experimenter would see their responses. The experimenter then started a computer program that provided participants with instructions and items for the questionnaires below. These questionnaires

were administered (in the order listed below) and responses were recorded using MediaLab software (Jarvis, 2000).

Dependent Measures

Immediate response. Participants’ typewritten responses to the confrontation were automatically recorded as part of the AOL Instant Messenger program.

Affect. A 32-item adjective checklist questionnaire assessed participants’ emotional responses to the experimental tasks and the confrontation. Participants were instructed that because their computer experience may be influenced by their mood, they were to indicate the extent to which they experienced each emotion at any point during the experiment. Participants rated each item (e.g., *happy*, *anxious*, *guilty*) separately on the computer using a 7-point scale ranging from 1 (*does not apply at all*) to 7 (*applies very much*).

Attitudes toward computers. To appear consistent with the cover story, participants completed a filler questionnaire about their experience using the computer during the two experimental tasks as well as general computer use.

Partner evaluation. The instructions to this questionnaire explained to participants that their attitudes toward computer use and the experimental tasks may have been influenced by their interaction with the other participant. Thus, 10 items assessed participants’ evaluation of their partner, (e.g., “I enjoyed working with my partner very much,” “I probably wouldn’t be friends with someone like my partner”).

Postconfrontation stereotypic responses. After completing the questionnaires on the computer program, the experimenter provided participants with several pages of photograph–sentence pairs similar to those used in the second computer task. The experimenter explained that it was important to examine how people respond to tasks they complete on a computer as well as on paper, so they were to provide handwritten responses to the photograph–sentence pairs. These photograph–sentence pairs were different from those used in the computer task but similar in nature. Of importance were participants’ responses to three postconfrontation critical photograph–sentence pairs. These pairs included different photographs and inferential statements but were similar to the prefrontation critical pairs in that both stereotype-consistent and stereotype-inconsistent responses were possible (e.g., This person is good at getting into locked doors: burglar or locksmith; This person takes cars from people: car thief or valet; This person uses needles for recreation: drug addict or tattoo artist). Seventeen filler items were also included in this task. Participants were instructed that their responses would be completely confidential and they were to place the completed questionnaire in a sealed receptacle.

After completing the questionnaires, participants were debriefed and probed for suspicion. All participants believed they were interacting with another participant. Because deception had been used, participants were given the opportunity to disallow their data to be included (all allowed their data), were given their research credit, thanked, and dismissed.

Results and Discussion

Participant Prejudice

The 20-item ATB showed strong internal consistency ($\alpha = .90$) and a composite prejudice score was computed. Although participant prejudice scores ranged from 1.00 to 6.47 (possible range: 1–7), the distribution was somewhat skewed toward less prejudiced scores ($M = 2.94$, $SD = 1.02$). A univariate analysis of variance (ANOVA), using gender and threat condition to predict ATB scores, revealed a marginally significant effect of gender such that men were marginally more prejudiced ($M = 3.11$) than women ($M = 2.78$), $F(1, 117) = 3.09$, $p = .08$. Participants’

prejudice scores did not differ as a function of experimental condition, $F(1, 117) = 0.17, p = .69$, and there was no interaction between participant gender and threat condition ($F < 1$).

Summary of Analytic Procedure

Except where otherwise noted, all dependent variables were predicted using multiple regression with confrontation condition (0 = low threat, 1 = high threat) and participant prejudice scores (centered and treated continuously) entered at the first step and their interaction entered at the second step. Initial analyses that also included participant gender as a predictor variable yielded no significant main effects or interactions. In addition to the 121 participants reported above, 2 participants failed to provide any stereotypic responses to the initial photograph-sentence task (thus rendering confrontation inappropriate) and their data were excluded.

Initial Stereotypic Responses

Participants' responses to the three critical photograph-sentence pairs, which had been recorded by the AOL Instant Messenger service, were evaluated for their stereotypic content. Responses were clearly stereotype consistent or not, so interrater reliability was high (agreement across three statements = 98%). After removing the 2 participants who did not provide any stereotypic responses, the 121 remaining participants provided at least one stereotypic response, 83.5% provided at least two stereotypic responses, and 45.5% provided all three ($M = 2.29$). It is important that the number of initial stereotypic responses did not significantly differ as a function of threat condition, $F(1, 118) = 0.05, p = .83$, and was not related to participants' prejudice level, $F(1, 118) = 1.61, p = .21$.

Immediate Reactions to Confrontation

The majority of participants (78.5%) provided a typewritten reaction to the confrontation. These reactions were examined for recurring themes. Six categories were identified that captured the vast majority of responses. One judge coded all reactions for the presence or absence of each of the six response categories. Then, a second judge coded all responses; the judges demonstrated

acceptable interrater reliability (using Scott's, 1955, formula, $\pi = .86$).

The six response categories and their frequencies by condition can be seen in Table 1. *Acknowledgment* was defined as any form of verbal concession or understanding of the confronter's point of view (e.g., "I know what you mean," "I understand"). A response was considered *bias recognized* if participants expressed awareness of or admitted the possibility that their responses could have been influenced by race and therefore biased (e.g., "Now I realize that I too make some judgments too quickly," "It's sad but they were the first thoughts that came to my head"). A response was considered an *apology* if the participant expressed any sort of regret (e.g., "I apologize," "I'm sorry"). A response was coded in the category *race not a factor* if participants claimed that race did not influence their responses (e.g., "Being Black had nothing to do with it," "I wasn't really looking at the pictures"). A response was categorized as a *denial* if the participant denied or rejected the suggestion that he or she was a prejudiced person (e.g., "I'm not a racist"). Last, a *hostile* response was one in which the participant conveyed anger, resentment, or condescension (e.g., "Don't be so sensitive"). Only five responses were ambiguous enough that they could not be coded into any of the six categories (e.g., "Not really but ok"). Of the remaining 26 participants (21.5%) who did not provide any immediate response, 5 were in the low-threat condition and 18 were in the high-threat condition.

As displayed in Table 1, although there was considerable variability in participants' reactions, the most frequent responses involved acknowledging the confrontation and/or explaining that one's previous stereotype-consistent responses were unrelated to racial factors. To examine these data more thoroughly, two second-order, general reaction categories were formed from the six specific categories in Table 1. An Acceptance reaction index was formed by summing the number of apology, acknowledgment, and bias-recognized reactions. A Rejection reaction index was formed by summing denial, race not a factor, and hostile reactions. These indexes were computed by summing the number of reactions because participants often provided more than one type of response to the confrontation in the same typewritten message ($M = 1.36$). However, results were identical when the Acceptance and Rejection factors were coded dichotomously (1 = response present, 0 = response not present).

Table 1
Frequencies and Percentages of Typewritten Reactions to Confrontation

| Category | Reaction | Experiment 1 | | Experiment 2 | | Experiment 3 | |
|------------|-------------------|--------------|------|--------------|------|--------------|------|
| | | n | % | n | % | n | % |
| Acceptance | Acknowledgment | 55 | 33.5 | 52 | 29.9 | 24 | 30.0 |
| | Bias recognized | 19 | 11.6 | 20 | 11.5 | 7 | 8.8 |
| | Apologize | 10 | 6.1 | 15 | 8.6 | 7 | 8.8 |
| Rejection | Race not a factor | 58 | 35.4 | 67 | 38.5 | 31 | 38.8 |
| | Denial | 19 | 11.6 | 19 | 10.9 | 11 | 13.8 |
| | Hostile | 3 | 1.8 | 1 | 0.6 | 0 | 0 |
| Total | | 164 | 100 | 174 | 100 | 80 | 100 |

Note. Many participants provided more than one reaction to confrontation; thus, the total number of reactions is greater than the number of confronted participants in each experiment. Because of rounding, some values may not total to 100.

Correlations among these indexes and other dependent measures are provided below the diagonal of Table 2. Supporting our hypothesis, participants provided more accepting immediate reactions to a low-threat confrontation than a high-threat confrontation, $r(119) = -.31, p < .001$. It is interesting that there was a negative correlation between rejecting and participant prejudice such that relatively low-prejudice participants were more likely to provide rejecting reactions than relatively high-prejudice participants. Perhaps because responding without prejudice is central to the self-concept of low-prejudice individuals, they were quick to distance themselves from the possibility of being prejudiced by rejecting the confrontation.

Affective Reactions to Confrontation

Ratings for the 32-item affect measure were submitted to a principal components factor analysis with varimax rotation and produced three main factors. The strongest factor, Negself, involved feelings of negative, self-directed affect and included the following 10 items: angry at myself, annoyed at myself, depressed, disappointed in myself, disgusted with myself, guilty, regretful, low, self-critical, and shameful ($\alpha = .94$). A second factor, Negother, included negative emotions that were directed toward the participant's partner: angry at my partner, bothered, disgusted with my partner, irritated at my partner, and frustrated ($\alpha = .89$). A third factor, Discomfort, captured feelings of general anxiety and uneasiness and included the following eight items: anxious, embarrassed, fearful, sad, tense, threatened, uncomfortable, and uneasy ($\alpha = .91$).

Negself. Surprisingly, there were no significant effects associated with Negself. Participants did not differ as a function of threat condition (low-threat $M = 2.17$, high-threat $M = 2.35$), $F(1, 118) = 0.52, p = .48, \beta = .07$. In addition, contrary to past research showing that low-prejudice individuals experience greater negative, self-directed affect than high-prejudice people when they become aware of their prejudiced responses (e.g., Devine et al., 1991), the relation between prejudice and Negself was not significant, $F(1, 118) = 0.02, p = .89, \beta = .01$. The interaction between threat condition and prejudice was not significant, $F(1, 117) = 0.07, p = .80$.

Negother and discomfort. Consistent with expectations, participants in the high-threat condition experienced more anger and irritation toward the confronter ($M = 3.41$) than those who encountered a less threatening confrontation ($M = 2.72$), $F(1, 118) = 4.84, p = .03, \beta = .20$. When predicting Discomfort, there was a marginally significant main effect of threat condition, $F(1, 118) = 3.50, p = .06, \beta = .17$. Participants in the high-threat condition ($M = 2.94$) felt slightly more uncomfortable than those in the low-threat condition ($M = 2.46$). There were no significant main effects or interactions with participant prejudice.

Confronter Evaluations

The 10 items that measured participants' attitudes toward their partner (the confronter) demonstrated good internal consistency ($\alpha = .87$) and were combined to form an overall confronter evaluation index (after reverse scoring when appropriate). There was a marginal effect of prejudice such that less prejudiced participants evaluated the confronter slightly more favorably ($M =$

Table 2
Correlations Among Reactions to Confrontation for Experiments 1 and 2

| Variable | Threat condition | Participant prejudice | Immediate acceptance | Immediate rejection | Negself affect | Negother affect | Discomfort affect | Confronter evaluation | Stereotypic response |
|-----------------------------|------------------|-----------------------|----------------------|---------------------|----------------|-----------------|-------------------|-----------------------|----------------------|
| Confronter race | — | .09 | .16 | .09 | .29* | .06 | .10 | .02 | -.03 |
| Participant prejudice | .03 | — | .06 | .06 | .01 | .09 | .08 | -.01 | -.04 |
| Immediate acceptance | -.31** | .02 | — | -.22* | .27** | -.22* | .04 | .17 | .11 |
| Immediate rejection | .13 | -.20* | -.20* | — | -.02 | .26* | -.04 | -.06 | .11 |
| Negself affect | .07 | .01 | .28** | .07 | — | .12 | .62** | -.03 | .29* |
| Negother affect | .20* | .05 | -.17 | .24* | .33** | — | .32** | -.60** | -.04 |
| Discomfort affect | .17 | .12 | .07 | -.01 | .72** | .56** | — | -.16 | .07 |
| Confronter evaluation | -.17 | -.17 | .25** | -.09 | -.07 | -.57** | -.28** | — | .06 |
| Stereotypic response change | .05 | -.18* | .08 | .05 | .20* | .01 | .08 | .01 | — |

Note. Correlations below the diagonal correspond to Experiment 1; those above correspond to Experiment 2 (confrontation condition only). Stereotypic response change scores = preconfrontation minus postconfrontation (i.e., higher values correspond to decreased stereotypic responding). Threat condition scale values are as follows: 0 = low threat, 1 = high threat. Confronter race scale values are as follows: 0 = White, 1 = Black. Negself = factor involving feelings of negative, self-directed affect; Negother = factor involving negative emotions directed toward the participant's partner. * $p < .05$. ** $p < .01$.

5.07) than more prejudiced participants ($M = 4.71$), $F(1, 118) = 3.53$, $p = .06$, $\beta = -.17$. In addition, consistent with our hypothesis, participants confronted in a less threatening manner tended to evaluate their partner more favorably ($M = 5.07$) than participants confronted in a more threatening and accusatory manner ($M = 4.72$), $F(1, 118) = 3.39$, $p = .07$, $\beta = -.16$.

Postconfrontation Stereotypic Responses

Participants' handwritten responses to the three critical photograph-sentence pairs were evaluated for their stereotypic content. Similar to the initial critical photograph-sentence pairs, all responses were clearly either stereotype consistent or not, so interrater reliability was very high (agreement across three statements = 99%). A within-participant analysis showed that the number of stereotypic responses provided significantly decreased after confrontation (preconfrontation $M = 2.29$, postconfrontation $M = 1.06$), $t(120) = 14.18$, $p < .001$. It is clear that the confrontation was effective in decreasing stereotypic responses.

Did the effects of the confrontation on subsequent stereotypic responding depend on participants' threat condition or prejudice level? When predicting participants' change in stereotypic responding (i.e., preconfrontation responses minus postconfrontation responses), neither the main effect for threat, $F(1, 118) = 0.34$, $p = .57$, $\beta = -.05$, nor the interaction between threat and prejudice were significant, $F(1, 117) = 0.26$, $p = .61$. That is, although participants in the high-threat condition were feeling irritated and evaluated the confronter less favorably, threat condition did not influence the overall effect of the confrontation on stereotypic responding. This suggests that even hostile confrontations may be effective in reducing subsequent stereotypic responding. In contrast to the effect of threat condition, the effect for prejudice was significant such that relatively low-prejudice participants showed a greater decrease in stereotypic responding ($M = 1.41$) than relatively high-prejudice participants ($M = 1.05$) after confrontation, $F(1, 118) = 4.20$, $p < .05$, $\beta = -.19$. Thus, although all participants decreased stereotypic responding, less prejudiced participants were more likely to change than more prejudiced participants. Similar regression analyses predicting number of Time 2 stereotypic responses while controlling for Time 1 responses yielded identical findings.

Relations Among Measures

Table 2 (below the diagonal) shows the correlations among the various measures collected throughout the experiment. Participants' immediate reactions to the confrontation were associated with different affective responses. Participants who had more accepting public reactions to the confrontation tended to experience greater negative self-directed affect and evaluate the confronter more favorably. In contrast, participants who reacted in a rejecting manner tended to feel more anger and irritation toward the confronter. It is interesting that participants' immediate reactions were not related to change in stereotypic responding.

When examining relations among affective measures, an interesting positive correlation between Negself and Negothen emerged, $r(119) = .33$, $p < .01$. In past research examining people's affective reactions when their prejudice-related discrepancies were made salient (e.g., Monteith, Devine, & Zuwerink,

1993), participants either responded with Negself (low-prejudice participants) or Negothen (high-prejudice participants) but not both. However, of particular importance, there was a significant positive correlation between the extent to which participants experienced Negself and their decrease in stereotypic responding, $r(119) = .20$, $p < .05$. Supplemental analyses indicated that this effect was not moderated by the type of confrontation or participant prejudice level. Thus, consistent with previous research on the role of Negself in prejudice reduction (e.g., Monteith, 1993), participants who experienced higher levels of negative self-directed affect were more likely to provide fewer stereotypic responses after being confronted.

Summary

The results of Experiment 1 suggest that people's reactions to being confronted about bias may partially depend on the way in which the confrontation is worded. Participants had more accepting immediate reactions and felt less anger and irritation in response to a low-threat appeal for fairness compared with a more threatening accusation of racism. This is consistent with previous research (e.g., Baumeister et al., 1996; Monteith et al., 1996) that has suggested that most people are willing to embrace norms of egalitarianism (i.e., a low-threat confrontation) but react negatively to having their self-images impugned (i.e., a high-threat confrontation). However, we should note the possibility that our confrontation messages had different implications beyond their level of threat. That is, the high-threat confrontation urged participants to be less prejudiced, whereas the low-threat confrontation asked participants to be more fair. This subtle difference may have invoked different self-regulatory foci (i.e., avoidance-related prevention focus vs. approach-related promotion focus) that may have influenced participants' reactions to the confrontation (Higgins, Shah, & Friedman, 1997).

Perhaps most important, our findings indicated that both types of confrontations had the behavioral consequence of curbing subsequent stereotypic responding. Even if participants did not like their partner and were angry with the person, they heeded the confronter's suggestion. In addition, relatively high-prejudice participants reduced their stereotypic responses, albeit less so than relatively low-prejudice participants. Furthermore, participants were more likely to decrease their stereotypic responding to the extent that the confrontation elicited feelings of negative, self-directed affect. However, the absence of a correlation between confrontation type and Negself rules out the possibility of Negself mediating the relation between confrontation and changes in stereotypic responding. We suspected that this lack of correlation resulted because all participants were confronted in Experiment 1. Indeed, the lack of a no-confrontation control condition opens our findings to the possibility that simply participating in the experiment may have led to decreased stereotypic responding (i.e., even if participants were not confronted).

Experiment 2

We conducted a second experiment to include a no-confrontation control condition, which was necessary to increase our certainty that the confrontation caused participants to reduce stereotype-consistent responses. In addition, Experiment 2 allowed

us to test competing hypotheses regarding the potential influence of the confronter's race on reactions to confrontation. In our previous research (Czopp & Monteith, 2003), people indicated they would experience more negative self-directed affect when they imagined confrontations from a White person than similar confrontations from a Black person. Theoretically, this may be because nontarget confrontations are surprising and increase message processing. If such findings generalize to actual confrontations, participants should feel more *Negself* after being confronted by a White person than a Black person. Alternatively, if self-presentation concerns become especially salient in actual confrontations, participants may feel more *Negself* after being confronted about a racial bias by a Black person than a White person. However, because a confrontation from either person is likely to increase the salience of personal and social norms against prejudice, both types of confrontations (i.e., from a target and nontarget) should lead to decreases in stereotypic responding.

Method

Participants

Two hundred White participants were initially recruited from the department's introductory psychology participant pool. Of these 200, 7 failed to provide any initial stereotypic responses (and thus could not be confronted) and 6 expressed suspicion or experienced Internet connection problems. The data from the 187 remaining participants (90 men, 97 women) were included for analysis. Recruitment procedures were identical to those used in Experiment 1.

Procedure

In a 2×2 between-participants factorial design (with participant prejudice as an additional factor that varied continuously), participants interacted with either a White confederate or a Black confederate and were either confronted by the confederate or not. The cover story and procedure were very similar to that of Experiment 1. As in Experiment 1, the confederate provided scripted responses throughout the session. During the first computer task, in which the participant and confederate exchanged personal background information, the confederate explicitly indicated his or her race (White or Black). Gender was not mentioned. During the second interactive task involving responses to photograph-sentence pairs, the confederate provided nonstereotypic responses to the critical items if the participant was in the confrontation condition. If the participant was in the control condition, the confederate matched the type of response provided by the participants (i.e., if the participant provided a stereotypic response, so did the confederate; if the participant provided a nonstereotypic response, so did the confederate).

In addition, before the second task, the experimenter informed participants in the confrontation condition that when the task was over there would be a brief feedback session and that they were to exchange any comments regarding their partner's performance during the second task. Participants in the confrontation condition received the following feedback, which was intended to be a moderately threatening combination of the two confrontations from Experiment 1:

i thought you typed pretty fast, but i noticed that for some of the pictures of Black people you said some stereotypical things like criminal, bum, and welfare. i mean, i guess that fits, but maybe that's being a little biased. don't you think we should really try to treat everyone equally?

Confronted participants were given an opportunity to respond to the confrontation. Participants in the no-confrontation condition did not have a

feedback session and thus were not confronted. Participants were informed that the interactive portion of the experiment was over and that all responses they made from that point on would be completely confidential. Participants then completed, using MediaLab (Jarvis, 2000), the same 32-item affect questionnaire and 10-item partner evaluation questionnaire as in Experiment 1.

Afterward, the experimenter explained that the study also concerned memory skills during computer use. Under the pretense of examining their memory of the interaction, participants responded to several questions related to information about their partner as exchanged in the first task (e.g., "What is your partner's major?"). This allowed us to verify whether participants correctly recalled the race of their partner (all participants did). Last, participants completed a second photograph-sentence inference task on the computer and were instructed to distinguish stimuli they had previously seen and completed in the first computer task from novel items. Again, under the guise of a memory test, participants were instructed that they would see various photograph-sentence stimuli. If the stimuli had been previously seen, participants were to provide the same response they provided the first time. If the stimuli were novel, participants were to provide new inferences. The photograph-sentence pairs were randomly presented on the computer screen, and participants typed in their responses. Ten stimuli were previously seen (none were the preconfrontation critical stimuli) and 10 were novel (including the 3 postconfrontation critical photograph-sentence pairs). By presenting this second photograph-sentence task in the context of a memory experiment, we hoped to minimize any potential influence of demand characteristics.

After completing all measures, participants were thoroughly debriefed and probed for suspicion. After debriefing, participants were given the opportunity to disallow their data to be included (all allowed their data), were given their research credit, thanked, and dismissed.

Results and Discussion

Participant Prejudice

The 20-item ATB showed strong internal consistency ($\alpha = .88$) and a composite prejudice score was computed. As in Experiment 1, there was a wide range of prejudice scores (1.00 to 5.60), and the distribution was skewed toward less prejudiced scores ($M = 2.89$, $SD = 0.91$). Participants' prejudice scores did not differ as a function of confrontation condition, confronter race, or participant gender, nor were there any interactions (all $F_s < 1$).

Summary of Analytic Procedure

Except where otherwise noted, all dependent variables were predicted using multiple regression with confrontation condition (0 = control, 1 = confrontation), confronter race (0 = White, 1 = Black), and participant prejudice scores (centered and treated continuously) entered at the first step, two-way interactions entered at the second set, and the three-way interaction entered at the third step. Again, initial analyses that also included participant gender as a predictor variable yielded no significant main effects or interactions.

Initial Stereotypic Responses

Participants' responses to the three critical photograph-sentence pairs were recorded by the AOL Instant Messenger service and evaluated for their stereotypic content. Responses were clearly stereotype consistent or not, so interrater reliability was high (agreement across three statements = 99%). All participants pro-

vided at least one stereotypic response, 88.8% provided at least two stereotypic responses, and 59.9% provided all three stereotypic responses ($M = 2.49$).

The number of stereotypic responses was not related to participant prejudice, $F(1, 182) = 0.61, p = .44, \beta = .05$. However, there was a main effect of confrontation condition such that participants in the confrontation condition provided fewer stereotypic responses ($M = 2.24$) than participants in the control condition ($M = 2.74$), $F(1, 182) = 26.94, p < .001$. This result was most likely due to the fact that in the control condition, the confederate's responses matched participants' responses (i.e., if the participant gave a stereotypic response, the confederate also gave a stereotypic response), but in the confrontation condition the confederate always provided nonstereotypic responses. Participants in the confrontation condition may have been influenced by the confederate's nonstereotypic responses and consequently provided fewer stereotypic responses. If this was the case, likelihood of stereotypic responding should decrease during the course of the photograph-sentence task. Indeed, although participants in the two confrontation conditions were equally likely to provide stereotypic responses for the first critical photograph-sentence pair (92% in the control condition and 88% in the confrontation condition), $\chi^2(1, N = 187) = 0.84, p = .46$, participants in the confrontation were less likely to provide stereotypic responses than participants in the control condition for the second (73% vs. 96%, respectively), $\chi^2(1, N = 187) = 20.28, p < .001$, and third (63% vs. 88%, respectively), $\chi^2(1, N = 187) = 15.63, p < .001$, critical photograph-sentence pairs. Given these findings, responses to the initial critical items need to be controlled when analyzing responses to the final set of critical items.

Immediate Reactions to Confrontation

Participants' typewritten responses to the confrontation were recorded by the AOL Instant Messenger program and were content analyzed on the basis of the six categories of reactions established in Experiment 1. One judge then coded all responses for the presence or absence of each of the six response categories. Then, a second judge coded all responses and demonstrated acceptable interrater reliability (using Scott's, 1955, formula, $\pi = .86$). Only 8 participants failed to provide any typewritten response to the confrontation (all of whom were in the White-confronter condition). Reaction frequencies are displayed in Table 1, and supporting the validity of our coding schemes, the percentages closely correspond to those obtained in Experiment 1. Three responses were unrelated to the confrontation and were not coded (e.g., "you typed pretty fast"). As in Experiment 1, an Acceptance reaction index was formed by summing apology, acknowledgment, and bias-recognized reactions, and a Rejection reaction index was formed by summing denial, race not a factor, and hostile reactions. Correlations between these indexes and other dependent measures are provided above the diagonal of Table 2.

Affective Reactions to Confrontation

As in Experiment 1, ratings for the 32-item affect measure were submitted to a principal components factor analysis with varimax rotation and produced three main factors. Again, the strongest factor was Negself and included the following 10

items: angry at myself, annoyed at myself, depressed, disappointed in myself, disgusted with myself, guilty, regretful, low, self-critical, and shameful ($\alpha = .94$). The second factor, Negother, included the following items: angry at my partner, bothered, disgusted with my partner, irritated at my partner, and frustrated ($\alpha = .87$). The third factor, Discomfort, included 8 items: anxious, embarrassed, fearful, sad, tense, threatened, uncomfortable, and uneasy ($\alpha = .79$).

Negself. As in Experiment 1, there were no significant effects associated with participant prejudice level when predicting Negself. There was a significant main effect for confrontation condition such that confronted participants experienced significantly more Negself ($M = 2.17$) than participants who were not confronted ($M = 1.42$), $F(1, 182) = 25.70, p < .001, \beta = .34$. There was also a main effect of confronter race, $F(1, 182) = 9.92, p < .01$, qualified by a Confrontation \times Race interaction, $F(1, 179) = 3.80, p = .053$. Follow-up analyses revealed that the nature of the interaction was such that in the no-confrontation control condition, participants did not differ in reported Negself as a function of confronter race (White $M = 1.33$, Black $M = 1.51$). However, participants supposedly confronted by a Black person reported significantly more Negself ($M = 2.56$) than those confronted by a White person ($M = 1.81$), $F(1, 179) = 13.03, p < .001, \beta = .35$. It is interesting that this finding is contrary to the results found by Czopp and Monteith (2003), where participants reported they would feel more negative self-directed affect when confronted by a nontarget group member than a target group member. Czopp and Monteith suggested that nontarget confrontations may evoke greater Negself as a result of enhanced message processing (cf. Petty et al., 2001). However, this interpretation assumes that the confrontation message contained strong arguments. If a nontarget's confrontation leads to greater processing of a weak message, this may open the possibility for greater justification of the prejudiced response and consequently less Negself from a nontarget than a target confronter. Although not central to the goals of the current research, the precise role of confrontation-message elaboration is an important topic for future research.

Negother and discomfort. There was a substantial effect of confrontation condition on the Negother index such that confronted participants expressed greater irritation than nonconfronted participants (confront $M = 2.39$, control $M = 1.20$), $F(1, 182) = 56.84, p < .001, \beta = .49$. Confronted participants ($M = 2.52$) also experienced more discomfort than nonconfronted participants ($M = 1.95$), $F(1, 182) = 11.49, p < .001, \beta = .263$. Again, there were no main effects or interactions with prejudice.

Confronter Evaluations

The 10 items that measured participants' attitudes toward their partner demonstrated good internal consistency ($\alpha = .89$) and were combined to form an overall partner evaluation index (after reverse scoring when appropriate). Although all participants rated the confederate favorably, consistent with our hypothesis, confronted participants evaluated their partner significantly less favorably ($M = 5.02$) than participants who were not confronted ($M = 5.82$), $F(1, 182) = 31.25, p < .001, \beta = -.38$. There were no significant findings (main effects or interactions) involving confronter race or participant prejudice.

Postconfrontation Stereotypic Responses

Participants' responses to the second set of critical photograph–sentence pairs were recorded by MediaLab (Jarvis, 2000) and evaluated for their stereotypic content. Similar to the initial pairs, all responses were clearly stereotype consistent or not, so interrater reliability was high (agreement across three statements = 99%). Data from 19 participants were not available because they failed to follow instructions properly.²

As in Experiment 1, a stereotypic-response change score was computed for each participant by subtracting the number of postconfrontation stereotypic responses from the number of preconfrontation stereotypic responses (i.e., higher scores correspond with greater decrease in stereotypic responding). Supporting the prediction that confrontation decreases the likelihood of future stereotypic responses, confronted participants were much more likely to change ($M = 1.15$) than nonconfronted participants ($M = 0.21$), $F(1, 163) = 47.33$, $p < .001$, $\beta = .47$. The main effect of confronter race was not significant, $F(1, 163) = 0.34$, $p = .56$, $\beta = .04$, nor was the effect of participant prejudice, $F(1, 163) = 0.29$, $p = .59$, $\beta = -.04$. There were no interactions among these factors. This suggests that although the Black confronter elicited more Neginself than the White confronter, both were equally effective in decreasing subsequent stereotypic responding (White $M = 1.18$, Black $M = 1.12$).

A similar covariate analysis predicting Time 2 stereotypic responses was conducted using multiple regression. Time 1 stereotypic responses were added to the first step of the regression analysis along with confrontation condition, confederate race, and participant prejudice and corresponding interactions were entered in subsequent steps (i.e., two-way, three-way, and four-way interactions). Time 1 stereotypic responses significantly predicted Time 2 responses, $F(1, 162) = 21.80$, $p < .001$, $\beta = .28$. Consistent with the above analysis using the pre–post change score, the only significant finding was a main effect of confrontation condition such that confronted participants provided fewer Time 2 stereotypic responses ($M = 1.10$) than nonconfronted participants ($M = 2.50$), $F(1, 162) = 83.08$, $p < .001$, $\beta = -.54$.

Relations Among Measures

Correlations among the various dependent measures for confronted participants are shown above the diagonal in Table 2. As in Experiment 1, participants' immediate reactions were differentially related to their affective responses: Accepting reactions were associated with feelings of Neginself, and rejecting reactions were associated with Negother. Also consistent with Experiment 1, immediate reactions were not related to participants' decrease in stereotypic responding.

Another important replication was that Neginself was positively correlated with decreased stereotypic responding among confronted participants, $r(85) = .29$, $p < .01$. Confronted individuals were particularly likely to provide fewer stereotypic responses to the extent that the confrontation elicited feelings of negative self-directed affect. As in Experiment 1, this effect was not moderated by participant prejudice level. Additional regression analyses within the confrontation condition that included Neginself and confronter race as predictors and stereotypic change as the dependent variable did not yield a significant interaction, suggesting that the

relation between Neginself and stereotypic change did not differ as a function of confronter race.

Given the significant effect of confrontation on the extent to which participants experienced Neginself and their decreased stereotypic responding, we tested whether the effect of the confrontation on subsequent stereotypic responding was mediated by negative self-directed affect. Previous research involving the effects of self-confrontation on subsequent regulation of prejudiced responses has suggested that the experience of such negative affect is a key component in self-regulatory prejudice reduction (e.g., Monteith, 1993). In agreement with Baron and Kenny (1986), confrontation was related to increased Neginself ($\beta = .34$, $p < .001$) and decreased stereotypic responding ($\beta = .47$, $p < .001$). When stereotypic responding was regressed simultaneously on confrontation and Neginself, the effect of confrontation decreased slightly ($\beta = .41$, $p < .001$) and the effect of Neginself remained significant ($\beta = .20$, $p < .01$). Sobel's (1982) test of the indirect effect of confrontation via Neginself was significant ($Z = 2.37$, $p < .05$), suggesting that Neginself is a partial, but significant, mediator of confrontation.

Summary

The results of Experiment 2 largely support the findings of Experiment 1 and suggest that interpersonal confrontations by Black and White confronters are similarly effective in curbing stereotypic responding. Compared with a control group, confronted participants were less likely to make stereotypic inferences. Furthermore, underscoring the importance of compunction-related affect in the prejudice-reduction processes, confrontations elicited Neginself, which in turn significantly predicted decreased stereotypic responding.

However, there were two methodological concerns that prompted a third investigation. First, although Experiment 2 included a no-confrontation control condition, it may not have appropriately mirrored the experimental condition. That is, there was an additional difference between conditions that was confounded with the presence or absence of a confrontation. Specifically, during the initial photograph–sentence task intended to elicit stereotypic responses from participants, the confederate in the confrontation condition always provided nonstereotypic responses to the critical items. In the control condition, the confederate matched the participant's responses, and because most participants provided stereotypic responses, so did the confederate. Thus, it is possible that simply receiving nonstereotypic responses from the confederate, rather than the confrontation itself, may have led to our results.

A second concern involved our measure of postconfrontation stereotypic responding. In both studies thus far, our sole measure was a photograph–sentence inference task identical to the one used to initially elicit stereotypic responding. Although we attempted to alter the context of the postconfrontation task (e.g., completing the

² Participants were instructed to determine whether they had previously seen each photograph–sentence stimulus. If the stimulus was the same as one previously used in the experiment, participants were to provide the same response they provided the first time. If the stimulus was novel, they were to provide an original response. These participants responded incorrectly by typing the actual words *same* or *different* rather than original or novel responses.

measure by hand in Experiment 1, as a test of memory in Experiment 2), we wanted to explore the generalizability of interpersonal confrontations to other types of prejudice-related responses.

Experiment 3

The goal for Experiment 3 was to examine more stringently the role of confrontations in curbing biased responding. First, we attempted to isolate the effect of confrontation by including a more precise control condition in which all aspects of the interaction were identical to the confrontation condition except the actual confrontation. Furthermore, we included a second control condition that involved a negative confrontation that was unrelated to racial bias. This allowed us to test whether any unpleasant interaction involving criticism leads to decreased stereotypic responding.

Last, to examine the generalizability of the confrontation effect, we included two types of postconfrontation measures in addition to measuring stereotypic inferences in the context of the photograph-inference task. Although we have demonstrated that confrontations effectively decrease stereotypic responding, we wanted to examine other types of group-based responding. A new measure assessed self-reported racial attitudes (i.e., on the ATB). Responding to explicit racial attitude items involves controlled processing, which is highly flexible (e.g., Devine, 1989). We reasoned that if a confrontation about race-related responding increased participants' motivation to try to control their subsequent prejudiced responses, the implications of the confrontational appeal should be salient when reporting their explicit racial attitudes.

The second measure new to Experiment 3 involved evaluating the humor of racial jokes. In past research (Monteith, 1993; Monteith & Voils, 1998), joke evaluation tasks have been used as a measure of relatively spontaneous prejudiced responses. This is because people often respond to jokes before thinking carefully; people report that they are prone to evaluating jokes that play on stereotypes more positively than their personal standards suggest is appropriate (Monteith, 1993). Thus, although confrontations may be effective in curbing prejudiced responding on tasks that resemble the initial confrontation context or that afford a great deal of cognitive control (i.e., the photograph-sentence task and the ATB, respectively), confrontations may be less influential on prejudice-related responses that are relatively more spontaneous.

Method

Participants

As in Experiments 1 and 2, potential participants completed the ATB as part of a mass screening session at the beginning of the semester. One hundred twenty-three participants were recruited from the department's introductory psychology participant pool. Approximately two thirds of participants ($n = 81$) volunteered to participate by signing themselves up through the department's research coordination Web site. Of these participants, 10 students were not present at the initial mass screening session and consequently did not have ATB scores; their data were excluded from analysis. In addition, to ensure a wide range of prejudice scores among our sample, the remaining one third of participants were recruited from the top third of the distribution of ATB scores. Potential participants were contacted by telephone using procedures identical to the previous studies. The final sample consisted of 111 White participants (18 men, 93 women).

Procedure

The cover story and procedure were largely similar to Experiments 1 and 2, with some notable differences. The confederate's race was made explicit and always specified as White during the exchange of personal information task. During the second interactive task requiring participants to make inferences on the basis of photographs and descriptions, the participant and the confederate alternated their responses to stimuli such that, unlike in the previous studies, both did not respond to each stimuli (e.g., the participant responded to the first item, then the confederate responded to the second item, and so on). For all conditions, the order of stimuli presentation was predetermined so that the participant always responded to the three critical photograph-sentence pairs and the confederate never responded to any critical photograph-sentence pairs. The confederate's responses were scripted and identical across all conditions.

As in Experiment 2, there was a brief feedback session after the photograph-sentence task in which the confrontations were made. Participants in the no-confrontation control condition received a brief message stating, "i thought you typed fast. good job." Participants in the racial confrontation condition received the following message:

i thought some of your answers seemed a little offensive. the Black guy wandering the streets could be a lost tourist and the Black woman could work for the government. people shouldn't use stereotypes, you know?

Participants in the nonracial confrontation condition received the following message:

i thought some of your answers seemed a little goofy. a traveler spends time in airports? a librarian has a lot of books? couldn't you think of anything better than that?

For both confrontation conditions, the content of the message was appropriately customized on the basis of the actual responses given by the participant during the task. All participants were given an opportunity to respond to the confederate's message. Participants were then informed that the interactive portion of the experiment was over and that all responses they made from that point on would be completely confidential.

Using MediaLab (Jarvis, 2000), participants then completed the 32-item affect questionnaire, the 10-item partner evaluation questionnaire, and the second photograph-sentence task. Similar to Experiment 2, the photograph-sentence task was framed as a test of memory during computer use and required participants to distinguish novel stimuli from previously viewed stimuli. Among these stimuli were three critical items. The joke evaluation task was then introduced using MediaLab. The instructions stated that people use computers as a means of sharing a variety of information, such as jokes, but that no research has examined how funny people think jokes are when told in person versus when read over the computer. Participants were then informed that they would read and evaluate jokes that were collected from various Internet and e-mail sources. The 14 jokes presented during this task were pretested and varied in funniness and tastefulness. Among the jokes were 3 racial jokes (e.g., "What's the most confusing holiday for Black kids? Father's day."). Participants rated each joke using a 0 (*not funny at all*) to 10 (*extremely funny*) scale.

Last, participants completed a measure of racial attitudes. They were informed that Internet surveys and questionnaires are becoming increasingly popular and that it is important to compare how people respond to online surveys versus traditional paper-and-pencil or telephone surveys. However, there was no mention of the mass screening session at the beginning of the semester nor any of the questionnaires that were included in that session. Instead, participants were told that the computer would randomly select one of several public opinion polls on topics such as television, politics, and drug use. After agreeing to continue, the computer waited several seconds while a progress bar was shown on screen as if the

computer was randomly selecting a survey. All participants completed the 20-item ATB (Brigham, 1993), making ratings on a 1 (*disagree strongly*) to 7 (*agree strongly*) scale. Afterward, participants were probed for suspicion, debriefed, and issued course credit. Two participants, not included in the 111 reported above, expressed suspicion about the true purpose of the experiment and their data were excluded.

Results and Discussion

Participant Prejudice

A composite prejudice score was computed that was based on the ATB that participants completed at the beginning of the semester as part of the mass screening session ($\alpha = .90$). As in Experiment 1, there was a wide range of prejudice scores (1.10 to 6.05), and the distribution was skewed toward less prejudiced scores ($M = 2.98$, $SD = 0.97$). Participants' prejudice scores did not differ as a function of confrontation condition $F(1, 105) = 0.64$, $p = .53$. There was a main effect of participant gender such that men reported higher prejudice scores ($M = 3.63$) than women ($M = 2.82$), $F(1, 105) = 15.41$, $p < .001$. The interaction was not significant.

Summary of Analytic Procedure

Except where otherwise noted, all dependent variables were predicted using multiple regression. Because there were three conditions, two dummy-coded variables were created: DC1 (0 = *racial confront*, 1 = *nonracial confront*, 0 = *control*) and DC2 (0 = *racial confront*, 0 = *nonracial confront*, 1 = *control*). When both variables are entered in the regression equation, DC1 is the contrast between the racial confrontation condition and the nonracial confrontation condition, and DC2 is the contrast between the racial confrontation condition and the no-confrontation control condition. These two dummy-coded variables and centered ATB scores were entered at the first step and two-way interactions at the second step. For each dependent variable, additional analyses were conducted to test the contrast between the nonracial confrontation condition and the control condition.

Initial Stereotypic Responses

Participants' responses to the three critical photograph-sentence pairs were recorded by the AOL Instant Messenger service and evaluated for their stereotypic content. Responses were clearly stereotype consistent or not, so interrater reliability was high (agreement = 99%). For the three critical stereotype-eliciting items, the vast majority (84%) provided all three stereotypic responses and all participants provided at least one stereotypic response ($M = 2.84$, possible range = 0-3). The number of stereotypic responses was not related to condition or participant prejudice.

Immediate Reactions to Confrontation

Participants' typewritten responses to the feedback session were again content analyzed on the basis of the six reaction categories established from Experiments 1 and 2. Two independent judges coded all responses among participants who were confronted about a racial bias for the presence or absence of each of the six response categories and demonstrated acceptable interrater agreement (97%). Reaction frequencies are displayed in Table 1 and are generally consistent with our previous findings supporting our coding schemes. As before, a second-order Acceptance reaction index was formed by summing apology, acknowledgment, and bias-recognized reactions, and a Rejection reaction index was formed by summing denial, race not a factor, and hostile reactions. Correlations between these indexes and other dependent measures are provided in Table 3.

Affective Reactions to Confrontation

A principal components factor analysis with varimax rotation was conducted on ratings for the 32-item affect measure and produced three main factors that included the same items as in Experiments 1 and 2. Composite variables were created for Negself ($\alpha = .90$), Negrother ($\alpha = .97$), and Discomfort ($\alpha = .85$). Means for each of these variables by condition are presented in

Table 3
Correlations Among Reactions to Confrontation for Experiment 3

| Variable | Condition | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|--------------------------------|-----------|-------|-------|------|-------|-------|-------|------|------|------|
| 1. Participant prejudice | -.02 | — | | | | | | | | |
| 2. Immediate acceptance | | -.17 | — | | | | | | | |
| 3. Immediate rejection | | .16 | -.41* | — | | | | | | |
| 4. Negself affect | .29* | -.27† | .09 | -.21 | — | | | | | |
| 5. Negrother affect | .77* | -.09 | -.41* | .28† | .03 | — | | | | |
| 6. Discomfort affect | .12 | -.20 | -.01 | -.11 | .71* | .28† | — | | | |
| 7. Confronter evaluation | -.81* | .15 | .32* | -.22 | -.10 | -.67* | -.27† | — | | |
| 8. Stereotypic response change | .53* | -.19 | -.10 | .20 | .36* | .06 | .33* | -.12 | — | |
| 9. Racist joke evaluation | .11 | .45* | -.25 | .03 | -.32* | -.12 | -.22 | .07 | -.26 | — |
| ATB score change | .25* | .36* | -.11 | .13 | .15 | -.07 | .23 | .17 | .15 | -.13 |

Note. Stereotypic response change scores = prefrontation minus postconfrontation (i.e., higher values correspond to decreased stereotypic responding). Condition scale values are as follows: 0 = *no-confrontation control condition*, 1 = *racial bias confrontation condition*. Negself = factor involving feelings of negative, self-directed affect; Negrother = factor involving negative emotions directed toward the participant's partner; ATB = Attitudes Toward Blacks scale (Brigham, 1993); ATB change scores = prefrontation minus postconfrontation (i.e., higher values correspond to decreased self-reported prejudice).

† $p < .10$. * $p < .05$.

Table 4. As expected, participants confronted about a racial bias reported feeling significantly more Negself than participants in the control condition, $F(1, 107) = 8.51, p < .05, \beta = .30$. Participants confronted about a racial bias also felt slightly worse than participants in the nonracial confrontation condition, though the effect was only marginally significant, $F(1, 107) = 2.91, p < .10, \beta = .17$. There were no differences in reported Negself between participants in the nonracial confrontation condition and the control condition, $F(1, 107) = 1.19, p = .27, \beta = .12$. There was also a main effect of participant prejudice such that across all conditions, low-prejudice participants experienced greater Negself ($M = 2.13$) than high-prejudice participants ($M = 1.75$), $F(1, 107) = 4.14, p < .05, \beta = -.19$. This main effect of prejudice was not qualified by any interactions with condition.

Also supporting our hypotheses, participants confronted about a racial bias reported feeling significantly more Negoother than participants in the control condition, $F(1, 107) = 83.54, p < .001, \beta = -.74$. Participants confronted about a racial bias also felt more Negoother than participants similarly criticized in the nonracial confrontation condition, $F(1, 107) = 8.26, p < .01, \beta = -.23$. Participants in the nonracial confrontation condition felt more Negoother than participants in the control condition, $F(1, 107) = 34.29, p < .001, \beta = .49$. Thus, a confrontation of racial bias elicited considerably more anger and irritation than an interaction without any confrontation and more anger and irritation than a similarly negative interpersonal situation that did not involve a charge of bias. There was also a main effect of participant prejudice such that across all conditions, low-prejudice participants experienced greater Negoother ($M = 3.05$) than high-prejudice participants ($M = 2.50$), $F(1, 107) = 4.02, p < .05, \beta = -.14$. There were no interactions involving participant prejudice. Last, there were no significant effects associated with the Discomfort index, suggesting that all participants felt similar levels of discomfort.

Confronter Evaluations

The 10-item partner evaluation measure demonstrated strong internal reliability ($\alpha = .95$) and participants' scores on this measure generally reflected their emotional reactions described above. As shown in Table 4, participants confronted about a racial

bias evaluated their partner significantly more unfavorably than participants in the control condition, $F(1, 107) = 108.98, p < .001, \beta = .78$. Participants confronted about a racial bias also rated the confronter more negatively than participants in the nonracial confrontation condition, $F(1, 107) = 6.16, p < .05, \beta = .19$. Participants in the nonracial confrontation condition rated the confronter more negatively than participants in the control condition, $F(1, 107) = 56.61, p < .001, \beta = .59$. Last, there was a main effect of participant prejudice such that low-prejudice participants evaluated the confederate less favorably ($M = 4.36$) than high-prejudice participants ($M = 4.81$), $F(1, 107) = 5.48, p < .05, \beta = .15$. As with the affect data, this main effect of prejudice was not qualified by any interactions.

Postconfrontation Responding

Stereotypic inferences. Participants' responses to the second set of critical photograph-sentence pairs were recorded by Media-Lab (Jarvis, 2000) and evaluated for their stereotypic content. Similar to the initial pairs, all responses were clearly stereotype consistent or not, so interrater reliability was high (agreement = 99%). Data from 11 participants were not available because they failed to follow instructions properly (see footnote 2).

Consistent with Experiments 1 and 2, a stereotypic-response change score was computed for each participant by subtracting the number of postconfrontation stereotypic inferences from the number of preconfrontation stereotypic inferences (i.e., higher scores corresponded with a greater decrease in stereotypic responding). Supporting the prediction that confrontation would decrease the likelihood of future stereotypic responding, participants confronted about a racial bias were much more likely to change ($M = 1.63$) than nonconfronted participants in the control condition ($M = 0.55$), $F(1, 96) = 39.87, p < .001, \beta = -.61$. Racially confronted participants were also more likely to change than participants in the nonracial confrontation condition ($M = 0.25$), $F(1, 96) = 24.76, p < .001, \beta = -.48$. There was no difference in stereotypic-response change between participants in the nonracial confrontation condition and the control condition, $F(1, 96) = 1.85, p = .18, \beta = .13$. There were no significant findings (main effects or interaction) involving participant prejudice. Similar analyses

Table 4
Experiment 3: Means and Standard Deviations of Reactions to Confrontation by Condition

| Variable | Racial confrontation | | Nonracial confrontation | | No confrontation | |
|-----------------------------|----------------------|-----------|-------------------------|-----------|--------------------|-----------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> |
| Negself | 2.28 _a | 1.26 | 1.90 _{ab} | 0.82 | 1.63 _b | 0.79 |
| Negoother | 4.11 _a | 1.80 | 3.14 _b | 1.85 | 1.07 _c | 0.20 |
| Discomfort | 3.26 | 1.35 | 3.19 | 1.32 | 2.93 | 1.34 |
| Confronter evaluation | 3.59 _a | 1.02 | 4.16 _b | 1.36 | 6.01 _c | 0.65 |
| Stereotypic response change | 1.63 _a | 1.11 | 0.25 _b | 0.76 | 0.55 _b | 0.75 |
| Racist joke rating | 3.87 | 2.79 | 4.62 | 2.61 | 4.59 | 2.72 |
| ATB change | 0.20 _a | 0.54 | -0.03 _b | 0.58 | -0.09 _b | 0.60 |

Note. Means with different subscripts differ at $p < .05$. Negself = factor involving feelings of negative, self-directed affect; Negoother = factor involving negative emotions directed toward the participant's partner; ATB = Attitudes Toward Blacks scale (Brigham, 1993).

predicting Time 2 stereotypic responses that included Time 1 responses as a covariate yielded identical findings.

Racial joke ratings. Participants' ratings of the funniness of the three racist jokes were combined into a racist joke rating index ($\alpha = .87$). Although the means reported in Table 3 indicate that participants confronted about a racial bias evaluated the racist jokes as less funny than participants in the two control conditions, the difference between the racial confrontation and no-confrontation conditions was not significant, $F(1, 107) = 1.59$, $p = .21$, $\beta = .12$, and the difference between the racial and nonracial confrontation conditions was only marginally significant, $F(1, 107) = 2.25$, $p = .14$, $\beta = .14$. In addition, there was no difference between participants in the nonracial confrontation condition and the control condition, $F(1, 107) = 0.01$, $p = .78$, $\beta = .03$. As might be expected, there was a substantial main effect of participant prejudice such that high-prejudice participants found the racial jokes to be significantly funnier ($M = 5.63$) than low-prejudice participants ($M = 3.05$), $F(1, 107) = 30.91$, $p < .001$, $\beta = .48$. The same analysis conducted on a composite score of participants' ratings of the nonracial jokes ($\alpha = .82$) yielded no significant findings indicating that reported funniness of the non-racial jokes did not differ as a function of confrontation condition or participant prejudice level.³

Prejudiced attitudes. Participants' ratings for the 20-item ATB scale completed during the experimental session demonstrated high reliability ($\alpha = .89$), and scores were significantly correlated with ATB scores from the beginning of the semester, $r(109) = .80$, $p < .001$. The regression analysis of postconfrontation ATB scores (with pre-confrontation ATB scores included in the equation) revealed that, consistent with our hypothesis that interpersonal confrontation can decrease prejudiced responding, participants confronted about a racial bias reported less prejudiced attitudes than participants in the control condition, $F(1, 107) = 7.08$, $p < .01$, $\beta = .17$. The comparison between the racial and nonracial confrontation conditions was also significant, $F(1, 107) = 4.00$, $p = .05$, $\beta = .13$. As expected, the difference between the nonracial confrontation condition and the control condition was not significant, $F(1, 107) = 0.32$, $p = .58$, $\beta = .06$. Furthermore, confrontation condition did not significantly interact with participants' pretest ATB scores, indicating that the reporting of less prejudiced attitudes as a function of confrontation condition did not depend on initial level of prejudice. We also approached these data using a one-way ANOVA, with confrontation conditions as the independent variable and a difference score subtracting postconfrontation ATB scores from pretest scores as the dependent variable, and obtained the same findings (the data are presented in this form in Table 4). Additional analyses conducted only within the racial confrontation condition revealed that the difference between pre- and postconfrontation ATB scores ($M = 0.20$) was significantly different from zero, $t(40) = 2.34$, $p < .05$.

Relations Among Variables

As in Experiments 1 and 2, we examined correlations among the various dependent measures for participants confronted about a racial bias, and results were consistent with our previous findings. Correlations among all dependent variables among participants in the racial bias confrontation condition are provided in Table 3. Immediate acceptance of the confrontation was negatively corre-

lated with *Negother*, such that participants angered by the confrontation were less likely to type an accepting reply to the confronter, $r(39) = -.42$, $p < .05$. In addition, participants who typed an accepting response to the racial bias confrontation evaluated the confronter slightly more favorably than those who did not type an accepting reaction, $r(39) = .31$, $p < .06$. Again, immediate reactions to the confrontation were not related to any postconfrontation response measures (i.e., stereotypic inferences, joke ratings, or ATB scores).

Most important, Experiment 3 again implicated the role of *Negself* in the prejudice-reduction process. Consistent with Experiments 1 and 2, among participants confronted about a racial bias, *Negself* was significantly related to decreased stereotypic responding, $r(34) = .37$, $p < .05$. Furthermore, among participants confronted about a racial bias, those who felt more negative self-directed affect also evaluated the racist jokes less favorably, $r(39) = -.34$, $p < .05$. Thus, participants who experienced more *Negself* in response to a confrontation of racial bias provided fewer stereotypic inferences and evaluated the racist jokes as less funny than participants who felt less *Negself*.

Given the significant relations involving *Negself* and postconfrontation responding, we again tested *Negself* as a potential mediating variable. In agreement with Baron and Kenny (1986), confrontation (vs. no confrontation control condition) was related to increased *Negself* ($\beta = .35$, $p < .01$) and decreased stereotypic responding ($\beta = .50$, $p < .001$). When stereotypic responding was regressed simultaneously on confrontation and *Negself*, the effect of confrontation decreased ($\beta = .39$, $p < .01$) and the effect of *Negself* remained significant ($\beta = .31$, $p < .01$). Sobel's (1982) test of the indirect effect of confrontation on decreased responding via *Negself* was significant ($Z = 2.02$, $p < .05$), suggesting that *Negself* is a partial, but significant, mediator of confrontation. Thus, consistent with Experiment 2 and previous research on the role of *Negself* in prejudice reduction (e.g., Monteith, 1993), less stereotypic responding among participants confronted about a racial bias was partially mediated by the extent to which the confrontation elicited negative self-directed affect.

General Discussion

Findings from these three experiments suggest that confronting other people about their stereotypic responses results in a variety of intrapersonal and interpersonal outcomes. Emotionally, confronted participants consistently reported experiencing negative affect directed toward themselves (e.g., guilt and self-criticism)

³To examine the relation between participant prejudice and racial joke evaluation more fully, we also conducted a mixed-model ANOVA with confrontation condition and dichotomized prejudice level (based on a median split) as between-subjects variables and joke type (racial or non-racial) as the within-subject variable. There was a significant main effect of joke type qualified by an interaction between joke type and prejudice level, $F(1, 105) = 16.77$, $p < .001$. As might be expected, high-prejudice participants rated the racial jokes ($M = 5.46$) just as funny as the nonracial jokes ($M = 5.70$), whereas low-prejudice participants rated the racial jokes ($M = 3.24$) considerably less funny than the nonracial jokes ($M = 5.22$). Consistent with the reported regression analysis, the lack of a three-way ANOVA interaction suggests that these effects of prejudice level on joke evaluation did not differ as a function of confrontation condition.

and toward the confronter (e.g., anger and irritation). Confronted participants also consistently evaluated the confronter less favorably than nonconfronted participants. Our findings from Experiment 1 indicated that the extent of such negative reactions was influenced by the tone of the confrontational message. A hostile and accusatory confrontation that labeled participants as *racists* elicited stronger feelings of anger than a calm appeal for fairness and equality. Furthermore, our findings implicated the source of the confrontation as an important factor in determining people's reactions. In Experiment 2, a confrontation of racial bias purportedly made by a Black person elicited greater feelings of Negself than a confrontation made by a White person, although both confronters elicited greater Negself than was experienced in a nonconfrontation control condition.

Despite these differences in affective reactions as a function of experimental conditions, our findings indicated that confrontations of racial bias successfully reduced the likelihood of biased responses in a later experimental task, and this was true regardless of the tone and source of the confrontation. In Experiment 1, a confrontation that represented a harsh accusation of racism was just as effective as a less threatening plea for fairness in reducing biased responding. In Experiment 2, the Black and White confronters were similarly effective in decreasing participants' stereotypic responding. Last, in addition to providing fewer stereotypic inferences, confronted participants in Experiment 3 reported less prejudiced attitudes compared with nonconfronted participants.

Another consistent finding that should be underscored involved the role of negative self-directed affect in producing a reduction in stereotypic responses. Across all three studies, participants who experienced greater Negself in response to a confrontation of bias provided fewer stereotypic responses in a subsequent situation. Furthermore, among participants in Experiment 3, greater postconfrontational Negself was also associated with lowered funniness ratings for racist jokes. Last, in Experiments 2 and 3, Negself was a partial but significant mediator of the effect of confrontation on decreased stereotypic responding. These findings are consistent with previous research examining self-confrontation, which has shown that people who become self-aware of their prejudiced responses attempt to regulate and reduce them because of the experience of negative self-directed affect (e.g., Monteith, 1993; Monteith et al., 2002). Thus, confrontations from others are likely to be more effective (immediately and perhaps in the long term) to the extent that compunctious feelings such as guilt and self-criticism are elicited. Individuals may attempt to customize their confrontations so as to induce the most Negself (e.g., by increasing empathic concerns and/or attributions of culpability). However, although Negself was a partial mediator of the effect of confrontation, there remains quite a bit of variance to be accounted for in future investigations. Other potential mediators that may be examined in future research include external motivation to respond without prejudice (Plant & Devine, 1998) and public self-consciousness (Fenigstein, Scheier, & Buss, 1975). Because of the interpersonal nature of some confrontations, people may be more likely to adhere to the confronter's message to the extent that they are particularly attuned to themselves in a public context and wish to avoid negative reprisals from others.

What Is a Target to Do?

Eberhardt and Fiske (1996) reviewed and considered the active role that targets might play in encouraging people to reduce their prejudice, such as inviting identification by emphasizing shared goals. The present research provides another perspective relevant to the question of what a target can do. Research focusing on targets of stereotyping and prejudice suggests that many targets desire to confront instances of biased treatment, but they are less likely to do so when social costs are high (Shelton & Stewart, 2004) and when they fear negative interpersonal outcomes (Kaiser & Miller, 2004; Swim & Hyers, 1999). The present findings substantiate targets' concerns. Relative to nonconfronters, confronters of bias were liked less and more anger and irritation were directed toward them. Nevertheless, we think the results of paramount importance are that participants who were confronted did subsequently reduce their level of stereotypic responding (Experiments 1–3) and their reports of explicit prejudiced attitudes (Experiment 3). Confrontations elicited change, and the participants changed regardless of the hostility expressed in the confrontation, who did the confronting, whether participants immediately accepted or rejected the confrontation, and the extent to which participants disliked and felt negative affect toward the confronter. Because potential confronters often may be aware of only the immediate negative interpersonal outcomes of confrontations, it may be helpful to know that the confronted behavior may change even if these outcomes are negative. Furthermore, potential confronters may be willing to endure unpleasant interpersonal reactions if the confrontation will be ultimately successful in changing future behavior. An important issue to examine in future research is whether such negative interpersonal outcomes persist, with the most encouraging possibility being that they dissipate while behavioral change persists.

Different Confrontation Contexts, Different Effects?

Two findings that emerged consistently in the present research conflict with findings from other confrontation research. First, whereas Experiment 2 revealed that a Black confronter elicited greater negative self-directed affect than a White confronter, Czopp and Monteith (2003) and Mark et al. (2005) also found that a White confronter was more persuasive than a Black confronter. Second, although we found few effects related to participants' prejudiced attitudes, participants in Czopp and Monteith's research reported less negative self-directed affect in reaction to confrontation, and Mark et al. found that high-prejudice participants were marginally ($p < .07$) less persuaded by a confrontation compared with low-prejudice participants.

These differences may be due to the different contexts in which confrontation was examined across the various investigations. The present research involved an actual confrontation by another person who was present and engaged in the situation, which can be termed an interpersonal confrontation context. Because the person making the confrontation is directly involved in the interaction, the confronted person is likely to focus primarily on the confronter and how the confronted behavior may have caused offense. Indeed, being confronted by a person constantly referred to as one's *partner* throughout the experiment may have augmented such interpersonal concerns. We think that in such interpersonal con-

frontation contexts, people are likely to feel the most self-dissatisfaction over having been the cause of offense when the confronter is a member of the target group (e.g., a Black person). Recent research by Winslow (2004) has suggested that White individuals are particularly attuned to potential imputations of prejudice by Black individuals in interpersonal settings. Furthermore, the interpersonal nature of the situation may minimize the possible role of people's prejudice-related attitudes in shaping their affective reactions and subsequent behavior.

In contrast, the confrontation contexts used by Czopp and Monteith (2003) and Mark et al. (2005) involved noninterpersonal confrontations where participants reacted to imagined confrontation situations and to a confrontational newspaper editorial, respectively. To be sure, cases of and opportunities for noninterpersonal confrontations occur in everyday life, and people's reactions to them may depend on different factors than in interpersonal confrontations. In a noninterpersonal confrontation, the confronter is not actually present or engaged in the situation so there is less focus on others. Consequently, reactions may be largely influenced by relevant intrapersonal factors such as one's own personal attitudes (e.g., racial prejudice) and beliefs about the confronter or the confronter's group (e.g., as complainers; see Kaiser & Miller, 2001). We believe this is a reasonable working explanation for why our previous research using noninterpersonal confrontations found that target group members elicited less *Negself* than non-target group members and that prejudice-related attitudes factored into people's reactions to such confrontations.

This reasoning should be tested in future research that varies the context of the confrontation (i.e., interpersonal vs. noninterpersonal) while keeping other aspects of the experimental situation constant. Until then, other explanations for differences between the present findings and other confrontation findings cannot be ruled out. For example, perhaps prejudice was unrelated to participants' degree of negative self-directed affect because of the nature of the confronted responses, which may not have been indicative of actual prejudice or racial bias. The specific critical photograph-sentence stimuli were chosen because they were likely to yield stereotype-consistent responses that could then provide a basis for confrontation. However, in reality the sentences used would have yielded similar responses whether they were paired with photographs of White or Black individuals (see pilot testing reported in Monteith et al., 2002). Perhaps if participants' confronted responses more closely represented actual group-based discriminatory behavior—quite a challenge for controlled laboratory manipulation—the effect of participants' prejudiced attitudes would have been more consistent.

What Really Changes?

In each of the three experiments reported herein, confrontation was effective in decreasing stereotypic responses made in a context similar to the one that elicited the initially confronted behavior. That is, compared with control conditions, participants who made stereotypic inferences about Black targets became less likely to make similar inferences after being confronted by the confederate. A skeptical reader may regard these findings as mere compliance. Although the consistent correlations and mediations involving *Negself* suggest that such changes were at least partially based on internal motivations, even situationally compliant behav-

ior may be beneficial. By simply eliminating public displays of offensive behaviors by confrontations, social environments for targets of prejudice may be significantly improved. Furthermore, people's changed behavior may contribute to establishing and maintaining a normative environment that discourages prejudiced responses and consequently decreases the prejudiced responding of others (Monteith et al., 1996).

In Experiment 3, the effectiveness of confrontation generalized beyond participants' stereotyping during the photograph-inference task to participants' self-reported prejudiced attitudes. Compared with control conditions, participants confronted about a racial bias showed a greater decrease in ATB scores. Although this generalization of confrontation effectiveness beyond stereotypic responding is encouraging, we are cautious when interpreting these findings. Are these results indicative of actual attitudinal change, such that stored evaluative associations in relation to Black persons were altered? We suspect that such an immediate and drastic consequence of the confrontations used in the present research is unlikely. Nevertheless, at minimum, expressions of prejudiced attitudes on a confidential measure were altered, and we consider this to be an important outcome.

The results of Experiment 3 indicated that confrontation did not reduce the favorableness of participants' racial joke evaluations. Although condition means were in the hypothesized directions, these differences did not reach conventional levels of statistical significance. One possible explanation is that people's reactions to jokes that play on stereotypes are often spontaneous and more favorable than people conclude is appropriate after additional reflection (Monteith, 1993; Monteith & Voils, 1998). Because reactions to jokes, compared with reports of explicit racial attitudes, may be more difficult to control, the influence of confrontations on subsequent biased responding may have been limited in our studies to those responses that were easier to control. Moreover, participants in Experiment 3 were confronted about responses unrelated to racial jokes. Confrontations may prove more effective for reducing spontaneously biased responding when the confrontation specifically addresses such responses (e.g., in this case, if the initial confrontation had been related to racial jokes).

Similarly, the effectiveness of confrontations in encouraging the subsequent control of prejudiced responses may also depend on the extent to which confrontations serve to establish cues for control (Monteith, 1993; Monteith et al., 2002). Previous research on the self-regulation of prejudiced responses has demonstrated that people pay attention to environmental stimuli in situations involving prejudice-related transgressions and build associations among these stimuli, their prejudiced responses, and the negative consequences of those responses (Monteith et al., 2002; Monteith & Mark, in press). This serves to establish cues that, in the future, can trigger more careful responding so that prejudiced responses and the negative consequences of those responses can be avoided. Theoretically, confrontations related to biased responding should also serve to establish cues for control. For example, a particular confronter (e.g., a coworker), a particular situation (e.g., in class), or a specific type of biased response (e.g., laughing at jokes) may become associated with the confrontation and its consequences and develop into cues for control. To the extent that future situations include the presence of these cues, people should attempt to control and inhibit their prejudiced responses.

When considering the generalizability of our findings, it is important to acknowledge that this research focused exclusively on how White people respond to confrontations related to racially biased behavior targeting African Americans. Although we believe that many of the underlying processes involved in confrontation situations (e.g., the mediating role of Negself) may function similarly in the context of prejudice toward other social groups, there are likely important factors to consider. Confrontations may be less effective in situations that involve expressions of prejudice that can be easily justified or when there is relatively weak personal or social pressure to suppress such expressions (Crandall & Eshleman, 2003). For example, confronting an individual who openly derogates child molesters (a relatively justifiable prejudice) or confronting a White supremacist at a Ku Klux Klan meeting (where suppression norms are weak) may be particularly ineffective. Nonetheless, we feel that our findings involving confrontations are relevant for the vast array of situational contexts in which expressions of prejudice are considered inappropriate.

Furthermore, it is important to recognize that the confrontation context used in the present research involved temporary online interactions between two strangers who never saw each other. Although online computer communication (even between strangers) is becoming increasingly popular, other confrontation situations may involve close acquaintances or family members and may take place during face-to-face exchanges. Such interpersonal confrontation situations introduce a variety of factors (e.g., familiarity, physical proximity, repeated contact) that may yield reactions similar to or different from the present research. First, because people often wish to preserve their relationships with others, people may take a confrontation from a friend quite seriously and consequently comply with the confrontation's message. Alternatively, some individuals may feel secure enough in their relationship with the confronter to discount the confrontation without concern over its potential effect on the relationship. Second, some close friends or family members may be perceived as particularly credible sources of information (because they are respected and trusted), and consequently they may be more effective confronters. Of course, a confrontation from a less credible source (e.g., an annoying coworker who complains about everything) may undermine a confrontation's effectiveness. Third, we have argued that the experimental context in the present research differs from previous research (e.g., Czopp & Monteith, 2003) because of the increased emphasis on interpersonal concerns regarding the confronter. Such concerns are likely to be augmented in face-to-face situations where the confronter is physically close and one has the opportunity to read and respond to a variety of nonverbal indicators of disapproval (e.g., a disappointed tone of voice, facial features). As a result, people may be particularly responsive to such public and visible confrontations. Last, confrontation situations involving individuals who frequently interact with each other may be especially effective in changing behavior. A confronted person may change his or her behavior to avoid the discomfort of similar interactions with the confronter in the future. In addition, because of the repeated contact between the individuals involved in the confrontation, the confronter may become an especially strong cue for control for the confronted individual and facilitate self-regulation of future responding.

Indeed, the potential long-term persistence of the influence of interpersonal confrontations across time is an important issue to

consider in future research. Confrontations may have more lasting effectiveness among low-prejudice people because their prejudiced responses violate important personal standards of nonprejudice (Devine et al., 1991). Plant and Devine (1998) have found that low-prejudice people are predominantly internally motivated to avoid prejudice, and such intrinsic motivation is associated with greater perseverance (Deci & Ryan, 1985). In contrast, high-prejudice people are largely motivated to avoid prejudice for external reasons (e.g., social pressure to be politically correct). Consequently, the effectiveness of confrontation among high-prejudice people may be more temporary and situation specific. Moreover, confrontations involving people who are primarily externally motivated to control prejudice may even have some negative consequences in the form of attitudinal and behavioral "backlash" (see Plant & Devine, 2001).

In the beginning of this article, we invoked the words and imagery of the Civil Rights Movement as an illustration of how confrontations made by an individual can lead to large-scale improvements for many and induce societal level change. The extent to which individual level confrontations have ripple effects and lead to concerted efforts to promote change can be examined empirically. For example, perhaps being confronted will increase a person's likelihood of subsequently confronting others' behaviors. Having been personally influenced through a confrontation, a person may feel a sense of responsibility and opportunity to have a similar impact on others. Furthermore, merely observing a confrontation without being directly involved may have vicarious effects that similarly decrease biased responding while increasing the likelihood of future confrontations. Because people are often reluctant to confront others (Kaiser & Miller, 2004; Swim & Hyers, 1999) but also experience self-dissatisfaction when they fail to do so (Shelton et al., 2006), observing a model person confront prejudice may be useful. Thus, future research is needed to examine how confrontations may have a ripple effect in the spreading of antiprejudice norms and attitudes.

Conclusion

The capacity for people to influence others for the good of humanity is a compelling testament of the power of the individual. Confrontations may offer opportunities for individuals who wish to effect change in their social environment rather than passively tolerate others' prejudice. Additional research is needed to better understand what types of confrontations are most effective when delivered by whom and for what sorts of individuals. However, the present research offers an optimistic contribution by suggesting that interpersonal confrontation appears to be an effective means of decreasing biased responding. As American novelist Margaret Halsey (1946) wrote, "One of the less dismaying aspects of race relations in the United States is that their improvement is not a matter of a few people having a great deal of courage. It is a matter of a great many people having just a little courage."

References

- Alicke, M. D., Braun, J. C., Glor, J. E., Klotz, M. L., Magee, J., Sederholm, H., & Siegel, E. (1992). Complaining behavior in social interaction. *Personality and Social Psychology Bulletin*, 18, 286-295.

- America Online. (2000). AOL Instant Messenger 4.0 [Computer software]. Dulles, VA: Author.
- Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 361–382). New York: Guilford Press.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.
- Baumeister, R. F., & Campbell, W. K. (1999). The intrinsic appeal of evil: Sadism, sensational thrills, and threatened egotism. *Personality and Social Psychology Review*, *3*, 210–221.
- Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review*, *103*, 5–33.
- Blanchard, F. A., Lilly, T., & Vaughn, L. A. (1991). Reducing the expression of racial prejudice. *Psychological Science*, *2*, 101–105.
- Brigham, J. C. (1993). College students' racial attitudes. *Journal of Applied Social Psychology*, *23*, 1933–1967.
- Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin*, *129*, 414–446.
- Czopp, A. M., & Monteith, M. J. (2003). Confronting prejudice (literally): Reactions to confrontations of racial and gender bias. *Personality and Social Psychology Bulletin*, *29*, 532–544.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.
- 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. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology*, *60*, 817–830.
- Dylan, B. (1963). Oxford Town. On *The Freewheelin' Bob Dylan* [CD]. New York: Sony. (Most recent issue 2004)
- Eberhardt, J. L., & Fiske, S. T. (1996). Motivating individuals to change: What is a target to do? In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 369–415). New York: Guilford Press.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027.
- Feagin, J. R. (1991). The continuing significance of race: AntiBlack discrimination in public places. *American Sociological Review*, *56*, 101–116.
- Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, *43*, 522–527.
- Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). San Diego, CA: Academic Press.
- Halsey, M. (1946). *Color blind: A White woman looks at the Negro*. New York: Simon & Schuster.
- Higgins, E. T., Shah, J. Y., & Friedman, R. (1997). Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology*, *72*, 515–525.
- Jarvis, B. G. (2000). MediaLab Research Software (Version 2000) [Computer software]. New York: Empirisoft.
- Kaiser, C. R., & Miller, C. T. (2001). Stop complaining! The social costs of making attributions to discrimination. *Personality and Social Psychology Bulletin*, *27*, 254–263.
- Kaiser, C. R., & Miller, C. T. (2004). A stress and coping perspective on confronting sexism. *Psychology of Women Quarterly*, *28*, 168–178.
- Mark, A. Y., Monteith, M. J., & Czopp, A. M. (2005). *Legitimately confronting or just complaining?: The role of group membership in reactions to confrontations of prejudicial bias*. Manuscript submitted for publication.
- Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice reduction efforts. *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.
- Monteith, M. J., Deneen, N. E., & Tooman, G. D. (1996). The effect of norm activation on the expression of opinions concerning gay men and blacks. *Basic and Applied Social Psychology*, *118*, 267–288.
- Monteith, M. J., Devine, P. G., & Zuwerink, J. R. (1993). Self-directed versus other-directed affect as a consequence of prejudice-related discrepancies. *Journal of Personality and Social Psychology*, *64*, 198–210.
- Monteith, M. J., & Mark, A. Y. (in press). Changing one's prejudiced ways: Awareness, affect, and self-regulation. *European Review of Social Psychology*.
- Monteith, M. J., & Voils, C. I. (1998). Proneness to prejudiced responses: Toward understanding the authenticity of self-reported discrepancies. *Journal of Personality and Social Psychology*, *75*, 901–916.
- Monteith, M. J., & Walters, G. L. (1998). Egalitarianism, moral obligation, and prejudice-related personal standards. *Personality and Social Psychology Bulletin*, *24*, 186–199.
- Petty, R. E., Fleming, M. A., Priester, J. R., & Feinstein, A. H. (2001). Individual versus group interest violation: Surprise as a determinant of argument scrutiny and persuasion. *Social Cognition*, *19*, 418–442.
- 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. (2001). Responses to other-imposed pro-Black pressure: Acceptance or backlash? *Journal of Experimental Social Psychology*, *37*, 486–501.
- Rodin, M. J., Price, J. M., Bryson, J. B., & Sanchez, F. J. (1990). Asymmetry in prejudice attribution. *Journal of Experimental Social Psychology*, *26*, 481–504.
- Rokeach, M. (1973). *The nature of human values*. New York: Free Press.
- Rokeach, M., & Cochrane, R. (1972). Self-confrontation and confrontation with another as determinants of long-term value change. *Journal of Applied Social Psychology*, *2*, 283–292.
- Rokeach, M., & McLellan, D. D. (1972). Feedback of information about the values and attitudes of self and others as determinants of long-term cognitive and behavioral change. *Journal of Applied Social Psychology*, *2*, 236–251.
- Scott, N. A. (1955). Reliability and content analysis: The case of nominal scale coding. *Public Opinion Quarterly*, *19*, 321–325.
- Shelton, J. N., Richeson, J. A., Salvatore, J., & Hill, D. M. (2006). Silence is not golden: Intrapersonal consequences of not confronting prejudice. In S. Levin & C. Van Laar (Eds.), *Social stigma and group inequality: Social psychological perspectives*. Mahwah, NJ: Erlbaum.
- Shelton, J. N., & Stewart, R. E. (2004). Confronting perpetrators of prejudice: The inhibitory effects of social cost. *Psychology of Women Quarterly*, *28*, 215–223.
- Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equations models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290–312). San Francisco: Jossey-Bass.

Swim, J. K., & Hyers, L. L. (1999). Excuse me—what did you just say?!: Women's public and private responses to sexist remarks. *Journal of Experimental Social Psychology, 35*, 68–88.

Swim, J. K., Hyers, L. L., Cohen, L. L., Fitzgerald, D. C., & Bylsma, W. H. (2003). African American college students' experiences with everyday racism: Characteristics of and responses to these incidents. *Journal of Black Psychology, 29*, 38–67.

Tangney, J. P. (1995). Shame and guilt in interpersonal relationships. In J. P. Tangney & K. W. Fischer (Eds.), *Self-conscious emotions: Shame, guilt, embarrassment, and pride* (pp. 114–139). New York: Guilford Press.

Tangney, J. P., Wagner, P., Fletcher, C., & Gramzow, R. (1992). Shamed into anger? The relation of shame and guilt to anger and self-reported aggression. *Journal of Personality and Social Psychology, 62*, 669–675.

Winslow, M. P. (2004). Reactions to the imputation of prejudice. *Basic and Applied Social Psychology, 26*, 289–297.

Received May 9, 2003
 Revision received August 15, 2005
 Accepted September 12, 2005 ■

ORDER FORM

Start my 2006 subscription to *Journal of Personality and Social Psychology*! ISSN: 0022-3514

_____ \$202.00, **APA MEMBER/AFFILIATE** _____
 _____ \$421.00, **INDIVIDUAL NONMEMBER** _____
 _____ \$1,249.00, **INSTITUTION** _____
In DC add 5.75% / In MD add 5% sales tax _____
TOTAL AMOUNT ENCLOSED \$ _____

Subscription orders must be prepaid. (Subscriptions are on a calendar year basis only.) Allow 4-6 weeks for delivery of the first issue. Call for international subscription rates.



AMERICAN
 PSYCHOLOGICAL
 ASSOCIATION

SEND THIS ORDER FORM TO:
 American Psychological Association
 Subscriptions
 750 First Street, NE
 Washington, DC 20002-4242

Or call **800-374-2721**, fax **202-336-5568**.
 TDD/TTY **202-336-6123**.
 For subscription information, e-mail:
subscriptions@apa.org

Send me a FREE Sample Issue

Check enclosed (make payable to APA)

Charge my: VISA MasterCard American Express

Cardholder Name _____

Card No. _____ Exp. Date _____

 Signature (Required for Charge)

BILLING ADDRESS:

Street _____

City _____ State _____ Zip _____

Daytime Phone _____

E-mail _____

MAIL TO:

Name _____

Address _____

 City _____ State _____ Zip _____

APA Member # _____ *PSPA16*