

HOW SELF-AFFIRMATIONS INFLUENCE THE MOTIVATION AND
BEHAVIORS ASSOCIATED WITH PREJUDICE REDUCTION

by

Kristen Morton Vance

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

(Psychology)

at the

UNIVERSITY OF WISCONSIN-MADISON

1998

© Copyright by Kristen M. Vance 1998

All Rights Reserved

To My Family,
Especially Alexandra,
Who Inspired this Research

ACKNOWLEDGEMENTS

While conducting this research, I relied heavily on the unwavering support of several people, who I thank deeply:

Patricia Devine, my dissertation advisor, provided invaluable theoretical and methodological ideas, as well as critical feedback. Perhaps more importantly, through her example, Trish encouraged me to develop a passion for conducting experimental social psychological research that carried me through this project and will motivate me to undertake others.

As members of my dissertation committee, Colleen Moore, Janet Hyde, Carolin Showers and Jane Piliavin provided thoughtful advice concerning theoretical, methodological, and statistical issues. This project, and my growth as a social psychologist, benefited greatly from their time and contributions.

Several motivated and talented undergraduate students helped in conducting this research: Sarah Andrasky, Adam Barsky, Kassia Conway, Matt Koeser, Bob Peterson and Dan Segan. I gratefully acknowledge these individuals for their hard work and dedication to this project.

My graduate student colleagues were a wonderful source of support and ideas, especially Kenn Barron, David Ebenbach, Rosalie Margolis, Shawnee Parsil, Ashby Plant, Sarah Pressly and Kris Vasquez.

Invaluable financial support for this project came from two sources: an American Psychological Association dissertation award, and an award from the Department of Psychology Royalty Research Fund.

Finally, and most importantly, I thank Ezra, Mom and Alexandra, who loved, supported and encouraged me throughout. I am infinitely grateful for the personal sacrifices they endured so that I could pursue this degree.

TABLE OF CONTENTS

Dedication	i
Acknowledgements	ii
Abstract	vi
Introduction	1
The role of guilt in the prejudice reduction process	2
Alternative means of reducing discrepancy-related distress	5
Can people alleviate guilt through self-affirmation?	6
Goals of the present research	11
Method	
Design	16
Participants	17
Procedure	18
Dependent Variables	23
Results	30
Affective consequences	31
Behavioral consequences	36
Mediation of prejudice-relevant behavior	43
Discussion	47
Affective consequences	48

Behavioral consequences	52
Mediation of prejudice-relevant behavior	56
Alternative explanations	58
Theoretical implications and conclusions	63
References	69
Footnotes	73
Tables:	
1. Affect and confidence ratings in three pre-testing conditions: Baseline (no discrepancy), standards-only (essay) and discrepancy (essay and examples)	75
2. Pre-testing data for the four black jokes: Should ratings, would ratings, and discrepancy scores	76
3. Affect and confidence ratings in the control (no discrepancy), no self-affirmation, unrelated self-affirmation, and related self-affirmation conditions	77
4. Joke ratings, nonverbal reactions, and number of phone calls in the control (no discrepancy), no self-affirmation, unrelated self-affirmation, and related self-affirmation conditions	78
Appendices:	
A. Participant selection measures	79
B. Discrepancy manipulation	82
C. Self-affirmation materials	91
D. Affect and confidence materials	94
E. Joke rating task	99
F. Behavioral intention (phone call) measure	103

ABSTRACT

HOW SELF-AFFIRMATIONS INFLUENCE THE MOTIVATION AND
BEHAVIORS ASSOCIATED WITH PREJUDICE REDUCTION

Kristen Morton Vance

Under the supervision of Professor Patricia G. Devine

At the University of Wisconsin-Madison

Previous research has shown that when low-prejudice people behave in a manner inconsistent with their nonprejudiced standards, they feel guilty about the discrepancy and engage in a self-regulatory process that helps to inhibit future prejudiced responses. The present research tested the hypothesis that self-affirmations disrupt the prejudice reduction process by alleviating discrepancy-related guilt--the motivating force behind prejudice reduction efforts. Following activation of prejudice-relevant discrepancies, participants were randomly assigned to one of six experimental conditions in the "type of self-affirmation" (none, unrelated, and related) x "time of behavioral assessment" (immediate and two-day delay) factorial design, which included an appended no-discrepancy control condition. Participants' affect and confidence were measured immediately after discrepancy activation and self-affirmation. Prejudice-relevant behavior was measured either immediately after the self-affirmation or two days later by having

participants rate how funny they found four jokes that played on the stereotypes of black people. Self-reported ratings of joke humor as well as participants' nonverbal responses to the jokes were measured. Consistent with previous research, discrepancy activation produced high levels of guilt and decreased levels of positive affect for participants without self-affirmation opportunities. Further, important self-affirmations, both related and unrelated to the discrepancy domain, alleviated guilt and boosted positive affect following prejudice-relevant discrepancies. Surprisingly, participants who did not self-affirm reported higher levels of confidence than participants who self-affirmed. To test whether self-affirmations ultimately disrupted prejudice reduction efforts, prejudice-relevant behavior was examined. Results showed that participants who did not self-affirm following the discrepancy manipulation, those who were experiencing high levels of guilt, reacted less favorably to racial jokes than participants in a no-discrepancy control condition, who were not experiencing discrepancy-related guilt. In contrast, participants in the self-affirmation conditions, for whom guilt was alleviated, reacted more favorably to the jokes than no self-affirmation participants, and at levels comparable to control participants. This pattern of findings was obtained at both the immediate and two-day delayed assessments. Implications for the prejudice reduction process and self-affirmation theory are discussed.

Introduction

Despite widespread implementation of diversity training programs, recent incidents of sexual and racial discrimination in the military and large corporations suggest that an understanding of the variables that hinder and promote prejudice reduction is as important today as it ever has been (Reibstein, 1996; Vistica, 1996; Newman, 1997). Devine (1989) offered a contemporary theory of prejudice that reflects her optimism about the possibility for reducing prejudiced attitudes and behavior. This theory provides a useful starting point for examining prejudice reduction variables. Devine proposed that if people consciously choose to inhibit their stereotypes, they can learn to respond more consistently with their nonprejudiced standards. Achieving control over prejudiced responses after one has adopted nonprejudiced standards appears to be a gradual process rather than an all-or-none event, much like learning to "break a bad habit" (Devine, 1989; Devine, Monteith, Zuwerink, & Elliot, 1991).

Because the process occurs over time, learning to behave in a nonprejudiced manner is likely to be characterized by internal conflict when people occasionally "slip up" and respond in a more prejudiced manner than their personal standards permit. Devine and colleagues tested this notion by asking people how they thought they "should" respond to either Blacks or homosexuals in four hypothetical contact situations, and how they actually "would" respond in those situations (Devine et al., 1991; Monteith, Devine, & Zuwerink, 1993; Zuwerink, Devine, Monteith, & Cook, 1996). They found that low-prejudice people

who violated their personal nonprejudiced standards (that is, their "would" responses were more prejudiced than their "should" responses) experienced heightened levels of negative self-directed affect, specifically guilt, and decreased positive affect compared to low-prejudice participants without discrepancies and all high-prejudice participants. Further, personal standards regarding prejudice were found to be important, internalized and self-defining for low-prejudice people (Devine et al., 1991; Monteith et al., 1993; Zuwerink et al., 1996; Devine, Evett, & Vasquez-Suson, 1996). Thus, when low-prejudice people violate their internalized egalitarian standards, their nonprejudiced self-images are threatened and they feel guilty. An important question follows: How will low-prejudice people deal with the guilt they experience following situations in which they realize they have violated their egalitarian standards?

The Role of Guilt in The Prejudice Reduction Process

Devine and Monteith (1993) suggested that the guilt low-prejudice people experience when they violate their nonprejudiced standards motivates them to learn to inhibit future prejudiced responses (see also Monteith, 1993; Devine, 1995). Essentially, when low-prejudice people come in contact with minority group members, their stereotypes of that group will be automatically activated. If they respond in a manner consistent with their stereotypes, and if they evaluate their behavior as inconsistent with their nonprejudiced standards (behavior is more prejudiced than personal standards permit), low-prejudice people will experience guilt. At this point, a self-regulatory cycle is activated that helps low-

prejudice people learn to associate prejudiced responding with guilt, and should eventually help them to react in a more controlled, standard-consistent manner in the future.

To test this self-regulatory process model, Monteith (1993) conducted a study in which low- and high-prejudice participants in a discrepancy-activated condition were led to believe they had evaluated a gay law school applicant negatively because of his sexual orientation. Results indicated that when low-prejudice participants thought they had violated their internalized nonprejudiced standards by unfairly rating the gay law school applicant, they felt guilty, experienced heightened self-focus, were uniquely preoccupied with their personal prejudice-related discrepancy experiences, and attended carefully to discrepancy-relevant information, more so than low-prejudice participants who did not experience a discrepancy and all high-prejudice participants (Monteith, 1993, Study 1). These results provided evidence that low-prejudice people engage in a self-regulatory process following prejudice-relevant discrepancy experiences.

To examine the link between this self-regulatory process and subsequent prejudice-relevant behavior, Monteith (1993) conducted a second study in which prejudice-related discrepancies were activated in low- and high-prejudice participants via computer-based feedback which informed participants that their responses to ambiguous attitude items regarding homosexuals were "more prejudiced" than their prejudice-related standards indicated was acceptable. In a subsequent behavioral task, participants were asked to rate how funny they found

a series of jokes, including jokes at the expense of homosexuals. Low-prejudice participants who were presumably feeling guilty about the false feedback they just received rated the jokes less favorably than high-prejudice participants and other low-prejudice participants who were presumably not feeling guilty (Monteith, 1993, Study 2). Thus, prejudice-relevant discrepancies were shown to be related to activation of a self-regulatory process that helped low-prejudice participants successfully inhibit future prejudiced responding. Although Monteith demonstrated that guilt was a mediator of cognitive self-regulatory processes, her design did not permit a direct test of whether guilt mediated subsequent prejudice-relevant behavior. The mediational link between discrepancy-related guilt and behavior is, as yet, untested.

From a prejudice reduction standpoint, it is encouraging to know that when low-prejudice people "slip up" and act in a manner more prejudiced than their standards permit, guilt motivates them to self-regulate and thereby learn to avoid future prejudiced responses. However, learning to reduce prejudice is a demanding process which requires suffering unpleasant guilt and undergoing significant cognitive work both to detect the reasons for failing to respond according to nonprejudiced standards and to identify the cues that signal failure. Therefore, it is important to consider whether engaging in the laborious prejudice reduction process is the only means available to low-prejudice people for resolving guilt. If guilt can be alleviated through other means, it may not serve as a motivator and learning cue for prejudice reduction efforts.

Alternative Means of Reducing Discrepancy-related Distress

The cognitive dissonance literature suggests that discrepancy-related distress can be reduced in a variety of ways. Faced with an inconsistency between a belief (e.g., nonprejudiced standards) and a behavior (e.g., a prejudiced response), a person might change her or his behavior to be more in line with the personal belief (Festinger, 1957; Sherman & Gorkin, 1980). This behavior change strategy is compatible with Monteith's (1993) analysis. Alternatively, one might change his or her attitudes to be more consistent with current behavior. It is unlikely, however, that low-prejudice people would change their internalized, self-defining nonprejudiced attitudes in a more prejudiced direction. Research has shown that when an attitude is strong, central, important and self-relevant, attitude change is unlikely (Hardyck [Piliavin], 1966; Hardyck [Piliavin] & Kardush, 1968; Sherman & Gorkin, 1980).

A more plausible alternative for a low-prejudice person who has suffered a challenge to his or her nonprejudiced identity is to attempt to restore a positive self-concept (Aronson, 1968; Steele, 1988). Steele proposed that people desire to see themselves as "adaptively and morally adequate" (1988, p. 262), and that behaving inconsistently with one's attitudes threatens one's self-concept and motivates attempts to restore the self-image. Restoring one's self-image following a discrepancy experience may be accomplished in indirect ways, rather than by addressing and resolving the specific discrepancy. Steele (1988) provided an example of the smoker who wants to quit because he knows smoking is

hazardous to his health. The next time the smoker lights up, he might remind himself of what a good father he is in order to resolve the threat to his self-concept that the inconsistent smoking behavior poses. Self-affirmations, like the one the smoker used, are one way of restoring the self-concept through adding positive cognitions about the self. Thibodeau and Aronson (1992) characterized the self-affirmation process in this way: "After one is reminded of one's other positive qualities, the knowledge of a previous self-discrepant act becomes less relevant to one's self-definition and thus has far less power to threaten one's self-concept as a person of worth and integrity" (p. 598).

Can People Alleviate Guilt through Self-affirmation?

Because the prejudice reduction process is inherently lengthy, difficult, and conflict-laden, if other easier, more immediate opportunities to resolve guilt and restore a positive, nonprejudiced self-image were available prior to engaging in self-regulation, such self-affirmation opportunities might impede progress in prejudice reduction efforts (Devine & Monteith, 1993; Monteith, 1993). To examine this possibility, Vance, Devine and Barron (1997) tested whether three types of self-affirmations could effectively reduce the guilt that low-prejudice people experience following violations of their nonprejudiced standards. Discrepancies between participants' prejudice-relevant standards and their responses to homosexuals were activated using Devine et al.'s (1991) should-would task, a procedure that has previously been found to generate feelings of guilt among low-prejudice people with discrepancies (Devine et al., 1991;

Monteith et al., 1993; Zuwerink et al., 1996). After prejudice-related discrepancies were activated, participants were randomly assigned to one of three self-affirmation conditions: affirmation of an unimportant self domain; affirmation of an important self domain, but one that was unrelated to prejudice; and directly related affirmation of one's nonprejudiced identity, a self-domain that is also important to low-prejudice people (Devine et al., 1991). All of the affirmations were designed to confront participants with evidence of their competence, their "adaptive and moral adequacy", in the particular self-domain to which they were assigned. Following the self-affirmation opportunity, participants completed a 35-item affect questionnaire and indicated how confident they were in their abilities to respond consistently with their standards for how to treat homosexuals.

Vance et al.'s (1997) three self-affirmation opportunities varied along two dimensions in terms of their theoretical power to alleviate the distress that participants felt as a result of discrepant should-would responding: importance of the affirmation domain, and relatedness to the discrepancy domain. With regard to importance, Steele's (1988) research suggests that only affirmations of important self-domains should effectively reduce discrepancy-related distress. In terms of affirmation relatedness, the evidence is mixed regarding the question of whether related or unrelated affirmations should more effectively relieve distress. Some theorists suggest that the unique experience of guilt requires direct resolution in the specific domain that caused the guilt (Frijda, 1988; Tangney,

1995). Stone, Cooper, Weigand and Aronson (1997) provided empirical evidence that people prefer to directly resolve discrepancies, even if offered the chance to affirm a more important but unrelated self-domain. This position suggests that when low-prejudice people violate their nonprejudiced standards, they may be motivated to restore their nonprejudiced identities through direct, prejudice-related means.

In contrast, Thibodeau and Aronson (1992) proposed that focusing directly on the threat domain, even via self-affirmation, may actually heighten dissonance by reminding the person of a recent threat to the self-image. This viewpoint suggests that unrelated affirmations should reduce discrepancy-related distress more effectively than related affirmations because they minimize the relative importance of the recent inconsistent behavior in comparison to the overall positive self-concept. Aronson, Blanton, and Cooper (1995) found evidence that people prefer unrelated over related self-affirmations after experiencing discrepancies, and Blanton, Cooper, Skurnik and Aronson (1997) demonstrated that related self-affirmations actually increased attempts to reduce discrepancy-related distress through self-justificatory attitude change. This position suggests that low-prejudice people may be motivated to restore their nonprejudiced identities through unrelated means following prejudice-relevant discrepancies.

The extant empirical evidence is inconsistent regarding the efficacy of related versus unrelated self-affirmations. Vance et al.'s (1997) study, in addition to contributing to the analysis of prejudice reduction variables, was designed to

compare the efficacy of various types of self-affirmations for alleviating discrepancy-related distress. In terms of affirmation importance, within-condition analyses showed that only the important, unrelated affirmation effectively alleviated discrepancy-related guilt. The unimportant affirmation did little to relieve guilt. Low-prejudice participants with discrepancies in the unimportant affirmation condition felt more guilty than low-prejudice participants without discrepancies, replicating Devine et al.'s (1991) finding that when low-prejudice people experience prejudice-related discrepancies without a successful opportunity to self-affirm, they feel guilty (see also Monteith et al., 1993; Zuwerink et al., 1996; Monteith, 1993). Surprisingly, and contrary to Steele's (1988) prediction, the related nonprejudiced affirmation did not relieve guilt, despite the fact that it represented an important self-domain to low-prejudice participants.

With regard to affirmation relatedness, Vance et al.'s (1997) results showed that the unrelated affirmation alleviated guilt, but the related affirmation did not. This finding is most consistent with Thibodeau and Aronson's (1992) argument that by reminding the person of the recent threat to her or his self-image, related affirmations facilitate rather than alleviate feelings of discrepancy-related distress. In sum, the important, unrelated affirmation effectively alleviated guilt, presenting a potential threat to the prejudice reduction process. However, Vance et al.'s (1997) design did not permit a test of whether this threat actually disrupted subsequent prejudice reduction efforts.

In addition to analyzing the effects of self-affirmations on guilt, Vance et al. (1997) examined the influence of self-affirmations on positive affect and confidence in ability to act according to nonprejudiced standards, two other variables that theoretically might influence the motivation to reduce prejudice. Between-condition analyses revealed that low-prejudice participants with discrepancies in the unrelated affirmation condition and the related affirmation condition felt more positive than low-prejudice participants with discrepancies in the unimportant affirmation condition. These same participants also felt more confident in their ability to respond to homosexuals consistently with their standards. These findings suggest that both the unrelated and the related affirmations boosted positive affect and confidence for participants who experienced prejudice-relevant discrepancies.

In sum, low-prejudice participants who experienced prejudice-related discrepancies without an effective opportunity to self-affirm, those in the unimportant self-affirmation condition, felt guilty, negative and lacked confidence in their ability to respond without prejudice. Participants with discrepancies in the related self-affirmation condition felt guilty, but were also positive and confident. Finally, participants in the unrelated self-affirmation condition did not feel guilty; instead they felt positive and confident. Vance et al.'s (1997) results beg the question of which affective patterns (e.g., guilt combined with positive affect and confidence, versus guilt alone, versus positive affect and confidence alone) lead to future success at inhibiting prejudiced responses. Is guilt necessary for

motivating prejudice reduction efforts? If so, people who affirm themselves in an important, unrelated self domain, thereby alleviating guilt, may not inhibit future prejudiced responding. Or perhaps guilt is sufficient but not necessary to promote prejudice reduction, and people who feel positive and confident will be motivated to engage in the process without experiencing guilt. Even if guilt is necessary, it may not be sufficient for motivating prejudice reduction efforts. Perhaps the process is optimized for people who are motivated by guilt, but also feel positive and have confidence that they can achieve success at being nonprejudiced in the future. The explicit role of guilt in mediating prejudice-relevant behavior has not yet been fully tested.

Goals of the Present Research

Affective consequences of self-affirmation

One goal of the present research was to replicate Vance et al.'s (1997) findings regarding the affective consequences of self-affirmation using an alternative methodology. Vance et al. (1997) used a methodology for activating prejudice-related discrepancies that a) relied on having participants imagine themselves in various hypothetical scenarios, and b) resulted in a wide range of discrepancy magnitudes. That is, some participants displayed no inconsistencies between their "should" and "would" responses, while others displayed large discrepancies. The methodology in the present study required participants to draw on their own personal experiences interacting with minority group members,

and manipulated the presence or absence of prejudice-related discrepancies in the same way and with the same magnitude for all participants.

The experimental design in the present study tested the effects of three types of self-affirmation on prejudice-relevant discrepancy-related affect: no self-affirmation; important, unrelated self-affirmation; and related self-affirmation. It was predicted, based on Vance et al.'s (1997) data, that the unrelated self-affirmation would relieve guilt over levels associated with the no self-affirmation condition, but that the related self-affirmation would not relieve guilt. Both of the self-affirmations were expected to increase positive affect and confidence over levels experienced in the no self-affirmation condition. Thus, main effects for self-affirmation type were expected for guilt, positive affect and confidence. These analyses provided a test of the robustness of Vance et al.'s (1997) findings.

Behavioral consequences of self-affirmation

Vance et al. (1997) provided evidence that self-affirmations might pose a threat to prejudice reduction by showing that important, unrelated self-affirmations were effective at reducing guilt in low-prejudice people following violations of their nonprejudiced standards. What Vance et al. (1997) did not directly test was the link between the affective state following self-affirmation and subsequent prejudice-relevant behavior. Monteith (1993) demonstrated the connection between a violation of one's nonprejudiced standards and self-regulation processes, mediated by guilt, which were associated with inhibition of future

prejudiced responding. To convincingly argue that self-affirmations threaten this process, it is necessary to demonstrate that once guilt is relieved via self-affirmation, future prejudiced responding fails to be inhibited. The present study tested this possibility by adding a behavioral measure that reflected prejudice-relevant responding to Vance et al.'s (1997) design. A joke rating task was used to assess the behavioral consequences of self-affirmation following a prejudice-relevant discrepancy experience. Participants rated how funny they found a series of racial jokes; more favorable joke ratings reflected more prejudiced responses. In addition to the joke rating task, which reflected inhibition of prejudiced responding, a behavioral intention measure was included to assess participants' willingness to volunteer to actively promote fair treatment of minorities.

Predictions regarding prejudice-relevant behavior were based on Monteith's (1993) finding that low-prejudice people who experienced prejudice-relevant discrepancies, and were presumably feeling guilty, rated jokes at the expense of homosexuals as less funny than low-prejudice people for whom a prejudice-relevant discrepancy was not activated. A main effect for type of self-affirmation was predicted for joke ratings, such that participants experiencing the highest levels of guilt, those in the no self-affirmation condition, were expected to rate the jokes least favorably. It seemed plausible that a main effect for self-affirmation type would be present for the behavioral intention measure as well,

based on the expectation that participants experiencing the highest levels of guilt would be most motivated to volunteer.

Although the primary purpose of the present study was to determine whether the reduction of guilt would short-circuit the prejudice reduction process, another way to conceptualize the issue is in terms of what combination of affective states is most effective for facilitating behavior change. Perhaps feeling guilty about one's past behavior, combined with positive affect and confidence regarding future behavior, is a more powerful predictor and motivator of future nonprejudiced behavior than simply whether or not guilt is in force. A control theory analysis is consistent with this hypothesis, suggesting that positive affect and confidence in one's ability to reduce a discrepancy lead to reassertion attempts, which is reflected by increased efforts to do well on a subsequent task (Carver, Blaney, & Scheier, 1979; Carver & Scheier, 1990). Negative affect and low confidence that the discrepancy can be altered lead to an impetus to withdraw, reflected by reduced efforts at a subsequent task. The present study investigated whether positive affect and confidence, independent of guilt, predicted prejudice-relevant behavior following violations of nonprejudiced standards and various self-affirmation opportunities. Mediational analyses explored the roles of guilt, positive affect and confidence in mediating the prejudice reduction process.

Immediate versus long-term behavioral effects

In addition to assessing immediate prejudice-relevant behavior, as Monteith (1993) did, this study also measured long-term behavioral consequences by asking half the participants to complete the joke rating task and behavioral intention measure immediately after the discrepancy experience and self-affirmation opportunity. The other participants completed the behavioral measures two days after discrepancy activation and self-affirmation. From a practical standpoint, when a low-prejudice person violates his or her nonprejudiced standards and feels guilty, he or she may not encounter another prejudice-relevant situation for some time. It is important, therefore, to determine whether discrepancy-related guilt only motivates immediate behavioral inhibition, or persists to motivate longer-term prejudice reduction efforts.

Findings for immediate prejudice-relevant behavior were expected to replicate Monteith's (1993) results. Specifically, participants who experienced prejudice-relevant discrepancies, without the opportunity to self-affirm, were expected to react least favorably to the jokes. Whether this same pattern would hold for delayed assessments of prejudice-relevant behavior was an empirical question. Monteith (1993) suggested that if the prejudice-relevant discrepancy experience was strong enough to engage the self-regulatory process, inhibition of prejudiced responses may result even after the passage of time. However, Fried and Aronson (1995) found that discrepancy experiences improved recycling behavior immediately but not two weeks post-discrepancy. They speculated that

their participants were able to find other routes for reducing discrepancy-related distress in the “real world” after they left the lab, or that the distress generated by the discrepancy experience simply did not last. To address this issue, the present study examined the influences of guilt, positive affect and confidence on both immediate and long-term prejudice reduction efforts following prejudice-related discrepancies and various types of self-affirmations.

Method

Design

Experimental conditions

The design of this study is a three (type of self-affirmation: no affirmation; important, unrelated affirmation; and related affirmation) by two (time of behavioral measurement: immediate and two-day delay) factorial design, with one appended control group. Prejudice-related discrepancies were activated in the six experimental conditions by having low-prejudice participants engage in a prejudice-relevant discrepancy manipulation (described in detail in the “Procedure” section). Participants were then randomly assigned to one of the three self-affirmation conditions, which varied in terms of their theoretical power to reduce the guilt associated with prejudice-relevant discrepancies. Following the self-affirmation opportunity, affect and confidence were assessed. Finally, for half the participants, prejudice-relevant behavior was assessed immediately using a joke rating task and a behavioral intention measure. The other participants returned two days later to complete the behavioral measures, providing a more

long-term assessment of prejudice-relevant behavior following discrepancy experiences and self-affirmation.

No-discrepancy control condition

A no-discrepancy control condition was included in which participants first reported affect and confidence, then completed the behavioral measures, without experiencing the discrepancy manipulation (no prejudice-related discrepancies). This control condition provided a basis for comparing the affective and behavioral outcomes associated with each experimental condition relative to levels when no prejudice-related discrepancy was activated.

Participants

One hundred and twelve Introductory Psychology students at the University of Wisconsin were recruited for participation in this study, 16 participants per condition, in exchange for extra credit points.¹ Students were pre-screened on several variables of interest included in mass testing. Participants' prejudice toward Blacks was assessed using Brigham's (1993) Attitudes Toward Blacks Scale (ATB), included in Appendix A. Students who scored in the upper half of the ATB distribution, suggesting they were low in prejudice toward blacks, were considered potential participants (ATB score > 5.65 on a scale ranging from 1 to 7). Of this group, only students who self-reported being prone to discrepancies in responding to racial jokes were recruited. This was determined by asking mass survey participants to rate how often the following statements were true for them personally, on a one ("never" true) to seven ("always" true) scale: "I should find

jokes about Black people funny" and "I do find jokes about Black people funny". Only low-prejudice participants with positive should-would discrepancies ("do's" more prejudiced than "should's") were recruited. In all, 56% of mass survey respondents with ATB scores greater than 5.65 reported positive should-would discrepancies. This last qualification was necessary to insure that the behavioral measure, based on responses to racial jokes, was a valid indicator of prejudiced responding for all participants. That is, participants reported that they might laugh at racial jokes but believed they should not, hence finding humor in racial jokes presumably violated their nonprejudiced standards. Finally, because this study focused on prejudice toward blacks, African-American students were not recruited.

Procedure

Prejudice-related discrepancy manipulation

A modification of the "hypocrisy" paradigm devised by Stone, Aronson, Crain, Winslow, and Fried (1994) was used to activate prejudice-relevant discrepancies. Stone et al.'s (1994) hypocrisy manipulation consisted of asking participants to make a public statement about a firmly held belief (e.g., why it's important to practice safe sex), then asking participants to recall excuses they might have used when they failed to live up to that belief (e.g., excuses for why they failed to practice safe sex). Stone et al. (1994) found that when participants were both "committed" to a belief, as a result of publicly endorsing the belief, and "mindful" of times when they failed to exemplify the belief, they adjusted future

belief-relevant behavior to be more in line with the belief (e.g., by purchasing condoms), presumably because they felt hypocritical. This manipulation has been used to effectively instate hypocrisy (activate discrepancies) in the domains of recycling, water conservation behavior, and practicing safe sex (Fried & Aronson, 1995; Fried, in press; Dickerson, Thibodeau, Aronson, & Miller, 1992; Stone et al., 1997).

In the present study, participants were asked to write essays about why they felt it was important to treat black people fairly, then to recall two instances in which they failed to live up to this nonprejudiced standard. Appendix B includes the instructions and materials used in the discrepancy manipulation. Participants were contacted by a caller who said that he or she was scheduling participants for two unrelated studies, either back-to-back or two days apart depending on whether the participants were being assigned to the immediate or long-term condition. Participants were asked to commit to participating in both studies. When they arrived in the lab, participants were told that the purpose of the first study was to help the Student Orientation, Advising and Registration (SOAR) program coordinators strengthen the segment on diversity and human relations for the 1998 program. SOAR and its diversity component are actual programs that most participants took part in during the summer before their freshman year. Participants were told that because of recent incidents of discrimination on campus, the University was trying to improve its efforts to combat unfair treatment of minority students. The SOAR program was described as a good forum for

conveying this message to incoming students. Participants were told that young adults are more receptive to information regarding discrimination when it comes from peers as opposed to authority figures. Thus, the SOAR program was seeking peer inputs to its diversity and human relations segment.

To foster participants' commitment to their nonprejudiced beliefs, participants were asked to write brief essays about why they personally believed it was important to treat blacks on campus fairly. The essays would ostensibly be used by the SOAR program to help educate new students regarding diversity issues on the University of Wisconsin campus. Participants were told that portions of their essays would be published as part of a pamphlet, which would include portions of many students' essays about diversity issues and would be given to new students at SOAR 1998. After writing their essays, participants were asked to sign release forms authorizing SOAR to publish their essays with the participants' names attached.

To make participants mindful of times they violated their nonprejudiced standards, thereby activating prejudice-relevant discrepancies, participants were asked to write about two situations in which they reacted to black people more negatively than they thought was appropriate, or used or endorsed negative stereotypes of black people.

Pre-testing of the discrepancy manipulation. Although Stone et al. (1994) never directly measured affect following the hypocrisy manipulation, they assumed it generated feelings of distress, and specifically guilt. To verify whether

the prejudice-relevant discrepancy manipulation successfully generated guilt in low-prejudice people, the manipulation was pre-tested on 48 low-prejudice participants. Eighteen participants reported baseline levels of affect, in the absence of a prejudice-relevant discrepancy, as part of another study. Using the affect scale included in the present study, participants rated how much 35 affect items applied to them "at the moment" on a one to seven scale, where seven reflected strong endorsement of the affective state. Fourteen other low-prejudice participants completed the affect measure and confidence items used in the present study after writing essays about why they personally felt it was important to treat black people fairly. And 16 participants reported affect and confidence after both writing the essay and recalling two personal examples of times they had failed to live up to their standards for how to treat black people (the discrepancy manipulation used in the present study).

The results of discrepancy manipulation pre-testing are presented in Table 1. Results showed that levels of negative self-directed affect, guilt, were significantly different across conditions, $F(2, 45) = 10.13, p < .001$. A Bonferroni post hoc test revealed that levels of guilt in the discrepancy condition ($M = 3.47$) were higher than levels in both the essay-only and control conditions ($M = 1.75$ and $M = 1.81$ respectively), $p < .05$. Levels of global discomfort also differed significantly across conditions, $F(2, 45) = 5.75, p > .006$. Bonferroni post hoc indicated that levels of discomfort were higher in the discrepancy condition ($M = 3.37$) than they were in the control condition ($M = 2.04$), $p < .05$, with the mean for

discomfort in the essay-only condition not significantly different from the other two conditions ($M = 2.39$). Further, levels of positive affect differed significantly across the conditions, $F(2, 45) = 11.39, p < .001$. Bonferroni post hoc showed that levels of positive affect were lower in the discrepancy condition ($M = 2.83$) than in either the essay-only or control conditions ($M = 5.16$ and $M = 4.73$ respectively), $p < .05$. Means for depressed affect, negative other-directed affect, general confidence, and prejudice-relevant confidence did not differ significantly across conditions. In sum, participants in the discrepancy condition felt more guilty and uncomfortable as well as less positive than participants in the control and essay-only conditions. These data verify that the discrepancy manipulation used in the present study produced affective patterns that parallel those associated with other discrepancy manipulations that have been used in previous prejudice research (e.g., the should-would manipulation used by Devine et al., 1991; the false feedback manipulation used by Monteith, 1993), and therefore served the current purpose well.

Self-affirmation opportunity

After prejudice-related discrepancies were activated, participants in the six experimental conditions were randomly assigned to one of three self-affirmation conditions. The affirmation task was framed as another item of interest to the SOAR program coordinators, who were described as interested in learning more about “the other unique qualities of undergraduates at the University of Wisconsin”. In the no affirmation condition, participants were left to their own

devices for approximately five minutes (the average length of time required to complete the other affirmations), while the experimenter “went to get the survey they’re supposed to complete next”. The two self-affirmation conditions were adapted from Vance et al. (1997), and the affirmation materials are presented in Appendix C. In the related affirmation, participants were asked to generate four examples of times they were successful at “being nonprejudiced”. In the unrelated affirmation, participants were asked to generate four examples of times they were successful at some attribute they identified as very important to them personally. Vance et al. (1997) found that this idiographic affirmation task resulted in each participant affirming a self domain that was uniquely important to her or him.² Both of the affirmations were designed to confront participants with evidence of their competence, at least in one particular self-domain.

Dependent Variables

Affect measure

Following discrepancy activation and the self-affirmation opportunity, participants reported how they were feeling using the affect scale from Vance et al. (1997), shown in Appendix D. Participants were asked to indicate how they felt in the context of the study by rating how much 35 feeling words applied to them on a one (“does not apply at all”) to seven (“applies very much”) Likert scale. This affect measure permitted analyses of several specific affect indices: negative self-directed affect (guilt), global discomfort, positive affect, depressed affect, and anger at others.

Confidence measures

In addition to assessing how participants were feeling after the affirmation opportunity, the present study tested what effects, if any, the various affirmations had on participants' confidence regarding future nonprejudiced responding. Participants' confidence in their future ability to interact with blacks in a manner consistent with their nonprejudiced standards was assessed using the following item: "Over time, I am confident that I will be able to respond to blacks in a nonprejudiced manner." To test whether the various affirmations influenced confidence at a more general level, the following item, adapted from Rosenberg's (1965) self-esteem scale, was included: "Over time, I am confident that I will be able to do things as well as most other people."³ Participants responded to both items on a one ("not at all" confident) to fifteen ("extremely" confident) scale.

Behavioral measures

After affect and confidence were measured, half the participants in each affirmation condition immediately completed the behavioral measures as part of an "unrelated" study of humor. The other participants, who were already scheduled to return two days later for the "humor" study, were excused, and completed the behavioral measures during the follow-up session.

To examine whether type of self-affirmation and/or time of behavioral assessment influenced the inhibition of future prejudice-relevant behavior, participants were asked to report how funny they found a series of racial and sexist jokes. Disparaging humor often serves as a socially-acceptable way of

expressing hostility, dislike, or aversion that a person would be uncomfortable expressing more directly (Freud, 1928; McGhee, 1988; Neitz, 1980). Some people view racist and sexist jokes as innocuous, intended only to amuse. However, racist and sexist jokes can also be viewed as subtle indicators of prejudice because they often take advantage of racist and sexist stereotypes for their humor value, and endorse treating people in a prejudiced manner based on these stereotypes (LaFrance & Woodzicka, in press). Because participants in this study were selected based, in part, on their beliefs that they should not laugh at racist jokes, responding favorably to racist jokes was considered a form of prejudiced responding. Further, Monteith (1993) showed that low-prejudice participants who experienced prejudice-relevant discrepancies, and were presumably feeling guilty, rated jokes at the expense of homosexuals as less funny than low-prejudice participants for whom discrepancies had not been activated. Thus, in addition to representing a subtle behavioral indicator of prejudiced responding, the joke rating task has proven sensitive to differences in responding among low-prejudice people who were or were not experiencing prejudice-relevant discrepancies.

Pre-testing of joke stimuli. Pre-testing was conducted to identify racial jokes that were appropriate for use in this study. Participants were 27 Introductory Psychology students who scored low in prejudice on the ATB and reported proneness to laughing at racial jokes although they felt they should not (a sub-sample of the participant population selected for the actual study).

Participants were asked how funny they actually would find 36 racial and sexist jokes, on a one ("not at all funny") to eight ("extremely funny") scale. Eight of the jokes played on stereotypes of black people. Participants were also asked to rate how funny they should find the jokes, based on their personal prejudice-relevant standards. Order of presentation of the should and would ratings were counterbalanced. No significant order effects were found, so follow-up analyses examined joke ratings combined across order of presentation.

Based on analyses of these data, four of the black jokes and 14 of the non-black jokes were chosen for use in the present study. All of the jokes chosen were ones that participants said they should not find funny but did find moderately funny. That is, all the jokes selected for use in the study exhibited moderate humor ratings and were prone to discrepant responding. The results of pre-testing analyses on the four black jokes, which were of primary interest, are presented in Table 2. Joke 1 had a mean humor rating of 3.19, and a should mean of 2.00 (would - should discrepancy = 1.19). Joke 5 also had a mean humor rating of 3.19, with a mean should rating of 1.78 (would - should discrepancy = 1.41). Joke 10 had a mean humor rating of 4.19, with a mean should rating of 2.70 (would - should discrepancy = 1.49). And joke 15 had a mean humor rating of 3.11, with a mean should rating of 1.89 (would - should discrepancy = 1.22). Because all the target jokes included in the present study, including the non-black jokes, showed a positive discrepancy between should and would ratings of humor ("would's" were more prejudiced than "should's"), these

jokes afforded the possibility that humor ratings could be inhibited (decreased) if the self-regulatory process was engaged.

Inhibition of future prejudiced responses. The joke rating task was framed as an “unrelated study of humor”, and a second experimenter greeted each participant and conducted the session. To support the illusion of two unrelated studies, the discrepancy phase of the study was conducted in a seminar room in one wing of the psychology building, and utilized written stimulus materials. In contrast, the joke rating task was conducted in a lab room in the research wing of the same building and relied on computer presentation of the stimulus materials.

To determine whether participants were actively inhibiting prejudiced responses following the discrepancy experience and self-affirmation opportunity, participants were asked to rate how funny they found the 18 racial and sexist jokes identified in pre-testing. Ratings of the four black jokes constituted the primary dependent variable in this study. The instructions for the joke rating task and a list of the stimulus jokes are included in Appendix E. The first black joke, Joke 1, read: “How do you get 50 Haitians into a shoebox? Tell them it floats.” The second black joke, Joke 5, read: “Why is Stevie Wonder always smiling? He doesn’t know he’s black.” The third black joke, Joke 10, read: “Leroy is a 20 year old 9th grader who must use this word in a sentence: distress. ‘Distress . . . Girl, put on distress and let’s go, I be hungry’.” And the fourth black joke, Joke 15, read: “What’s the most confusing holiday for black kids? Father’s Day.” Participants rated how funny they found each joke on a one (“not at all funny”) to

eight (“extremely funny”) Likert scale, indicating their ratings in a response booklet.

In addition to assessing participants’ self-reported responses to the jokes, participants’ nonverbal emotional reactions were captured with a hidden video camera. Nonverbal reactions to the four target jokes served as a secondary dependent variable. During debriefing, participants were told they had been videotaped while they read and rated the jokes. They were asked for permission to use their tapes for research purposes. If they consented, participants signed release forms authorizing the use of their videotapes. If they declined, participants were offered the opportunity to have their tapes erased. No participant declined to allow the researchers to use the videotapes.

To avoid demand characteristics that might have biased the results of these measures, it was important that participants not realize the connection between the joke rating task and the previous discrepancy manipulation session. During debriefing, it was determined that only three participants guessed the connection between the two phases of the study, and their data were excluded from all analyses.

Willingness to volunteer. In addition to measuring whether participants inhibited personal prejudiced responding, participants’ willingness to volunteer to promote racial equality at a more societal level was assessed. This behavioral intention measure is presented in Appendix F. After announcing that the participant had completed all the tasks at hand and thanking her or him for

participating in the humor study, the experimenter asked each participant to complete one final form for SOAR which the previous experimenter "forgot" to give the participant earlier or at the previous session. The form advised participants that the SOAR coordinators were looking for volunteers to talk with new students, one-on-one, about the diversity information presented at SOAR. Participants were asked if they would be willing to make telephone calls to tell new students about the resources available on campus to help them treat African American and other minority students fairly. In addition to providing information, these phone calls would also give new students an opportunity to ask questions and discuss diversity issues in a more personal manner. The form allowed participants to check a box stating "No, I do not have time to volunteer" or "Yes, I would be willing to make _____ phone calls". Participants were told that calls should take about 5-10 minutes each. The instructions stressed that participants should only volunteer the number of calls they were genuinely willing to make. Participants were asked to write their local addresses and phone numbers if they agreed to volunteer so that the SOAR coordinators could contact them with details for making the phone calls.

To lower demand biases or feelings of obligation to the experimenter, participants were told that this volunteer activity was not directly related to the project they just participated in; the forms would simply be passed along to the SOAR coordinators. Fried and Aronson (1995) and Fried (in press) both found

this behavioral measure to be sensitive to differences in activated discrepancies in the domain of recycling.

Debriefing

After participants completed the behavioral measures, either during the immediate or delayed sessions, they were thanked for their participation and fully debriefed. The debriefing assured participants that the essays and examples they provided would be used for research purposes only, and would not be published by SOAR. Participants were also told that they would not be asked to fulfill any volunteer commitments they might have made. Participants were offered the opportunity to have their videotapes erased, or signed release forms if they consented to the use of their tapes for research purposes. Experimenters probed participants for suspicion, and asked if they had heard about the study prior to participating.⁴ Finally, if they were genuinely interested in volunteering to help SOAR with diversity efforts, participants added their names to a list to be given to the actual SOAR program coordinator at the conclusion of the study.

Results

Analyses were conducted to examine three central questions. First, to address the question of whether related and unrelated self-affirmations would alleviate discrepancy-related guilt, analyses tested the effects of self-affirmations on negative self-directed affect ("negself"). Second, to address the question of whether self-affirmations disrupted prejudice reduction efforts, the effects of self-affirmations on prejudice-relevant behavior (e.g., joke ratings and nonverbal

reactions) were investigated. Finally, mediational analyses were conducted to examine the mediational processes that governed the relationship between self-affirmations and behavior.

Affective Consequences of Discrepancy Activation and Self-affirmation

One important question in the present research was whether related and unrelated self-affirmations would alleviate discrepancy-related guilt. The experimental design permitted replication of previous findings concerning the affective consequences of prejudice-relevant discrepancies, in the absence of self-affirmation opportunities (Devine et al., 1991; Monteith et al., 1993; Zuwerink et al., 1996). In addition, the design permitted testing the potential guilt-relieving effects of related and unrelated self-affirmations (Vance et al., 1997). Of secondary interest was whether self-affirmations influenced positive affect and confidence.

Formation of affect indices

The 35 items that comprised the affect scale were used to construct five affect indices. Ratings of how guilty, angry at self, embarrassed, annoyed at self, disappointed at self, disgusted at self, regretful, ashamed, and self-critical participants felt were averaged to form a negative self-directed affect index, *negself* (Cronbach's alpha = .94). A global *discomfort* index included the items anxious, bothered, concerned, distressed, frustrated, negative, tense, uncomfortable, and uneasy (Cronbach's alpha = .89). A *positive* affect index was created by averaging the items friendly, happy, energetic, optimistic, content,

good, pleased with self, good about self, satisfied with self, and proud (Cronbach's alpha = .93). Ratings of how angry at others, disgusted at others, and irritated at others participants felt were averaged to form a negative other-directed affect index, *negother* (Cronbach's alpha = .91). Finally, a *depressed* index was created by averaging the items depressed, sad and helpless (Cronbach's alpha = .73). Previous research has established a precedent for using these indices to investigate the affective consequences of prejudice-relevant discrepancies (Devine et al., 1991; Monteith et al., 1993; Zuwerink et al., 1996; Vance et al., 1997).

Analyses of affect indices and confidence items

Each of the five affect indices and the two confidence items were examined in separate "self-affirmation type" (none, unrelated, and related) X "time" (immediate and delayed) between-subjects analyses of variance (ANOVA's). One-way ANOVA's with six non-orthogonal contrasts were also conducted to compare the experimental groups with each other and with the no-discrepancy control group. Because affect and confidence were measured before both the immediate and delayed behavioral assessments, affect and confidence were not expected to vary as a function of time. Therefore, no significant time effects were expected.⁵ Rather, *negself*, positive affect and confidence were expected to be linked to self-affirmation condition, and significant main effects for self-affirmation type were predicted. The data for the affect and confidence items are presented in Table 3.

Findings for negself. The two-way ANOVA for negself revealed the predicted significant main effect for self-affirmation type, $F(2, 89) = 7.39, p < .001$, with the no affirmation participants experiencing the highest levels of negself ($M = 4.15$), followed by the related affirmation participants ($M = 3.13$) and the unrelated affirmation participants ($M = 3.05$).⁶ There were no significant time or interaction effects for negself. All of the planned contrasts comparing conditions were significant, $p < .002$, except for the contrast between the related and unrelated affirmation conditions. These findings suggest that participants in the no affirmation conditions experienced higher levels of guilt than participants in the control condition ($M = 1.51$), replicating previous research that showed prejudice-relevant discrepancies, in the absence of self-affirmation, generate guilt in low-prejudice people (Devine et al., 1991; Monteith et al., 1993; Zuwerink et al., 1996; Vance et al., 1997). Further, both the related and unrelated affirmations alleviated guilt, relative to the no affirmation conditions, but did not completely relieve guilt back to levels reported by participants in the no-discrepancy control condition. In terms of alleviating guilt, these results suggest that both the related and unrelated affirmations were equally but not completely effective. These results are somewhat inconsistent with Vance et al.'s (1997) findings which showed that only the unrelated affirmation relieved guilt.

Findings for the other affect indices. The two-way ANOVA results for discomfort were similar to those for negself. The main effect for self-affirmation type was significant, $F(2, 89) = 4.30, p < .02$, such that the no affirmation

participants experienced the most discomfort ($M = 3.46$), followed by the related ($M = 2.92$) and the unrelated affirmation participants ($M = 2.69$). There were no significant time or interaction effects for discomfort. Because the findings for guilt and discomfort are so similar, and the two indices are highly correlated, subsequent analyses focused on guilt, the more theoretically relevant of the two indices.

The two-way ANOVA for positive affect showed the predicted significant main effect for self-affirmation type, $F(2, 89) = 5.70$, $p < .005$, with the no affirmation participants experiencing the lowest levels of positive affect ($M = 3.21$), followed by the related ($M = 3.81$) and unrelated affirmation participants ($M = 4.16$). There were no significant time or interaction effects for positive affect. All of the contrasts comparing conditions were significant, $p < .04$, except for the comparison of the related and unrelated affirmation conditions. These findings suggest that participants in the no affirmation conditions experienced lower levels of positive affect than participants in the control condition ($M = 4.97$), replicating previous research that showed prejudice-relevant discrepancies, in the absence of self-affirmation, decrease positive affect in low-prejudice people (Devine et al., 1991; Monteith et al., 1993; Zuwerink et al., 1996; Vance et al., 1997). Further, both the related and unrelated affirmations boosted positive affect, relative to the no affirmation conditions, but did not completely restore positive affect back to levels reported by participants in the no-discrepancy control condition. In terms of restoring positive affect, these results suggest that both the related and unrelated

affirmations were equally but not completely effective. These results are consistent with Vance et al.'s (1997) findings which showed that both related and unrelated affirmations increased positive affect.

The ANOVA's for negother and depressed affect showed no significant condition, time, or interaction effects. Consequently, these indices were excluded from subsequent analyses.

Findings for confidence. The two-way ANOVA for general confidence indicated the predicted significant main effect of self-affirmation type, $F(2, 89) = 3.69$, $p < .03$, but in an unexpected direction, such that participants in the no affirmation conditions reported the highest levels of general confidence ($M = 13.13$), followed by participants in the unrelated ($M = 11.84$) and related affirmation conditions ($M = 11.41$). There were no significant time or interaction effects for general confidence. Because general confidence was not measured for the control group, follow-up analyses were conducted by submitting general confidence ratings to a one-way ANOVA with three levels, combining each of the affirmation conditions across time. As expected, general confidence was found to vary across type of self-affirmation, $F(2,93) = 4.08$, $p < .02$. The Bonferroni post hoc test, $p < .05$, revealed that the no affirmation participants reported higher levels of general confidence than the related affirmation conditions. Levels of general confidence in the unrelated affirmation conditions did not differ significantly from the other conditions.

The two-way ANOVA for prejudice-related confidence showed no significant condition effects or interactions. Descriptive analyses indicated that this measure suffered from a ceiling effect. Means across conditions ranged from 13.25 ($sd = 1.57$) to 14.19 ($sd = .83$) on a one to 15 scale. It is likely that the restricted range of responses to this confidence item, as well as high standard deviations in the immediate assessment conditions, constrained the analyses testing condition effects.

Behavioral Consequences of Discrepancy Activation and Self-affirmation

The most important question in the present research was whether the self-regulatory prejudice reduction process, presumably activated by guilt associated with prejudice-relevant discrepancies, would be disrupted when guilt was relieved via self-affirmation. To examine this question, analyses were conducted to investigate whether participants would 1) inhibit stereotype-consistent responses to jokes about black people, 2) suppress their nonverbal reactions to racist humor, and 3) volunteer to make phone calls to facilitate diversity efforts.

Formation of the joke rating index

Participants' ratings of the four individual black jokes were averaged to create a combined joke index (Cronbach's $\alpha = .66$). Higher scores on the index reflected funnier joke ratings (more prejudiced responses). Analyses focused on reactions to the jokes as an index, rather than examining reactions to each of the four black jokes separately, for several reasons. First, a principal components factor analysis with oblimin rotation showed that the four jokes

comprised a single factor which accounted for 47.7 percent of the variance in scores. The scree plot of eigenvalues also supported a single-factor solution (Tabachnik & Fidell, 1996). Second, Guilford (1954) noted that "for research purposes, one can tolerate much lower reliabilities than one can for practical purposes of diagnosis . . . even though it may be of the order of only .50" (p. 388), especially when a scale with lower reliability is still useful for prediction, as this index proved to be. Guilford also suggested that when attempting to predict a criterion that is very complex factorially, as is the case with measuring prejudiced responses, bringing in additional common factors may increase construct validity despite lowering internal-consistency reliability. Finally, from a practical standpoint, the goal of the present research was to examine whether responses to black jokes in general are influenced by the independent variables, rather than drawing conclusions from reactions to specific jokes about black people.

Analyses of joke ratings

Ratings in the combined joke index were submitted to a "self-affirmation type" (none, unrelated, and related) X "time" (immediate and delayed) between-subjects analysis of variance (ANOVA). A one-way ANOVA with six non-orthogonal contrasts was also conducted to compare the experimental groups with each other and with the no-discrepancy control group. Monteith (1993) found that low-prejudice participants who experienced prejudice-relevant discrepancies, without an opportunity to self-affirm, rated jokes at the expense of homosexuals as less funny than low-prejudice participants for whom discrepancies were not

activated. Therefore, a main effect for type of self-affirmation was predicted, based on the expectation that participants in the no affirmation conditions would inhibit prejudiced responses the most and would rate the jokes least favorably, while control participants, who did not experience a discrepancy, were predicted to rate the jokes most favorably. The data for the joke index are presented in Table 4.

The two-way ANOVA for the combined joke index revealed a marginally significant main effect for self-affirmation type, $F(2, 89) = 2.46, p < .09$. As predicted, the no affirmation participants rated the jokes least funny ($M = 2.77$), followed by the unrelated affirmation participants ($M = 3.11$), and the related affirmation participants ($M = 3.54$). There were no significant time or interaction effects. The one-way ANOVA showed that the contrast between the no affirmation conditions and the control condition was marginally significant, $t(105) = 1.74, p < .09$, such that the no affirmation participants rated the jokes less funny than the no-discrepancy control participants ($M = 3.53$). Further, the contrast between the no affirmation conditions and the related affirmation conditions was significant, $t(105) = -2.15, p < .04$, such that the no affirmation participants rated the jokes less funny than the related affirmation participants. None of the other contrasts comparing conditions were significant. These results suggest that only participants in the no affirmation conditions significantly inhibited joke ratings on the combined index. The findings for the combined joke index suggest that both of the self-affirmations, and particularly the related self-affirmation, disrupted the

process whereby prejudiced responses are normally inhibited for low-prejudice people following prejudice-relevant discrepancies.⁷

No significant results for type of self-affirmation or time were found in the analyses of the 14 non-black jokes. This finding indicates that the activation of prejudice-relevant discrepancies based on reactions to black people uniquely affected these low-prejudice participants' ratings of black-specific jokes.

Analyses of nonverbal reactions to racial jokes

Nonverbal responses to the jokes were recorded for 104 of the 112 participants. Data were lost for eight participants due to experimenter error and video camera malfunctions. Two coders observed the videotapes and rated participants' responses to the black jokes. Four response categories were used to reflect participants' reactions to the jokes: blank expression, positive reaction, negative reaction, or mixed (positive and negative) reaction. Out of 416 joke reactions (4 black jokes x 108 participants), the two coders disagreed 65 times, yielding an overall agreement rate of 84%. A third coder resolved all discrepancies between the initial two coders.

Index of nonverbal reactions. Because descriptive analyses revealed that negative and mixed reactions occurred with low frequency, these two types of reactions were combined with blank expressions to form a "non-positive reaction" category. Thus, analyses focused on participants' nonverbal reactions as a dichotomous variable: positive reactions to the jokes (assigned a value of "1") versus non-positive (blank, negative, or mixed) reactions to the jokes (assigned a

value of "0"). An index of nonverbal reactions to the black jokes was created by summing participants' reactions to the four individual jokes. Scores on this index ranged from zero (no positive reactions to the jokes) to four (positive reactions to all four black jokes).

Findings for the nonverbal index. Scores on the nonverbal index were submitted to a "self-affirmation type" (none, unrelated, and related) X "time" (immediate and delayed) between-subjects analysis of variance (ANOVA). A one-way ANOVA with six non-orthogonal contrasts was also conducted to compare the experimental groups with each other and with the no-discrepancy control group. As was the case for the self-reported joke ratings, a main effect for type of self-affirmation was predicted, based on the expectation that participants in the no affirmation conditions would react least favorably to the jokes, while control participants were expected to react most favorably to the jokes. The data for the nonverbal index are presented in Table 4.

The two-way ANOVA for the nonverbal index revealed a significant main effect for self-affirmation type, $F(2, 84) = 3.08, p < .05$. The no affirmation participants showed the fewest positive reactions to the jokes ($M = .75$), followed by the unrelated affirmation participants ($M = 1.19$), and the related affirmation participants ($M = 1.40$). There were no significant time or interaction effects. The one-way ANOVA showed that the contrast between the no affirmation conditions and the related affirmation conditions was significant, $t(97) = -2.23, p < .03$, such that the no affirmation participants reacted less favorably to the jokes than the

related affirmation participants. Although it was not statistically significant, there was a trend for the no affirmation participants to react less favorably to the jokes than the unrelated affirmation participants as well, $t(97) = -1.48$, $p = .14$. None of the other contrasts comparing conditions were significant. These results suggest that only participants in the no affirmation conditions significantly inhibited positive nonverbal reactions to the black jokes. Findings for the nonverbal index support the general pattern of results for the self-reported joke ratings: no affirmation participants reacted the most negatively overall to the black jokes whereas related affirmation participants reacted the most favorably.

Multivariate analyses of joke reactions

In addition to considering the effects of the independent variables on the joke rating index and the nonverbal reaction index separately, these dependent variables were entered simultaneously into a "self-affirmation type" (none, unrelated, and related) X "time" (immediate and delayed) multivariate analysis of variance (MANOVA). This analysis revealed a marginally significant main effect for type of self-affirmation, $F(4, 168) = 1.96$, $p < .10$. The univariate test for the joke rating index was marginally significant, $F(2, 84) = 2.68$, $p < .07$, and the univariate test for the nonverbal index was significant, $F(2, 84) = 3.08$, $p < .05$. The multivariate tests for time and the interaction of self-affirmation type by time were not significant. The results of this MANOVA provide further evidence that self-affirmation condition influenced subsequent self-reported and nonverbal reactions to racial jokes.

Analysis of volunteer phone calls

The number of phone calls participants volunteered was submitted to a "self-affirmation type" (none, unrelated, and related) X "time" (immediate and delayed) between-subjects analysis of variance (ANOVA). Results for this measure, where high scores reflected greater willingness to volunteer to promote racial equality, were expected to be opposite the pattern for the joke ratings and nonverbal reactions, where high scores reflected more prejudiced responses. A main effect for type of self-affirmation was predicted, based on the expectation that participants in both of the affirmation conditions, who had their guilt alleviated, would be less motivated to promote racial equality and would volunteer fewer phone calls than participants in the no affirmation conditions, who continued to experience high levels of guilt. Data for the phone call measure are presented in Table 4.

The two-way ANOVA for number of calls revealed no significant effects for self-affirmation type, time, or the interaction of these two independent variables. Descriptive analyses revealed that responses to this measure fell in a restricted range, with the mean number of calls in each condition ranging from .60 ($sd = 2.41$) to 2.14 ($sd = 3.98$). It is possible that a floor effect on this variable constrained the analyses.

In addition to analyzing the number of phone calls volunteered, a Chi-squared test was conducted on the overall percentage of participants in each experimental condition who volunteered to make at least one call versus those

who did not volunteer. As was the case for number of phone calls, this analysis failed to show significant differences between the conditions, $\chi^2(2, N = 96) = 2.85$, $p = .24$.⁸ A floor effect was again evident, with only 17% of participants across conditions volunteering to make any calls.

Mediation of Prejudice-relevant Behavior

Mediational analyses were conducted to examine the role of guilt in motivating prejudice reduction efforts. Monteith (1993) showed that discrepancy-related guilt mediated cognitive self-regulation efforts, which theoretically facilitate the inhibition of prejudiced responses. The present analysis tested whether guilt directly mediated prejudice-relevant behavior. In addition to guilt, it was hypothesized that positive affect and confidence may play mediational roles in reducing prejudiced responses. This assertion was examined as well.

Evidence that affect/confidence mediated prejudice-relevant behavior required three findings in regression analyses (Baron & Kenny, 1986). First, type of self-affirmation must significantly predict joke ratings and nonverbal reactions, the behavioral outcome variables. Second, type of self-affirmation must significantly predict affect/confidence, the proposed mediating variables. Finally, affect/confidence must significantly predict joke reactions, and the significant relationships between type of self-affirmation and joke reactions must be attenuated or eliminated when affect/confidence is included in the regression model.

To test these necessary links in the mediational chain, a series of regression models were estimated, in accordance with Baron and Kenny's (1986) guidelines. First, joke ratings and nonverbal reactions were regressed (in separate regression equations) on type of self-affirmation. Next, affect and confidence were regressed separately on type of self-affirmation. Finally, joke ratings and nonverbal reactions were regressed separately on both affect/confidence and type of self-affirmation at the same step in the regression model. Because type of self-affirmation is a categorical variable with three levels, two orthogonal contrast codes were created to represent this independent variable in the regression models. Contrast code one (CC1) compared the no affirmation conditions to the average of the self-affirmation conditions. Contrast code two (CC2) compared the related to the unrelated self-affirmation conditions. Both contrast codes were entered together as a block to represent type of self-affirmation in the regression models. Because the previous analyses showed no significant main effects or interaction effects for time of assessment, mediational analyses considered the effects of the three self-affirmation conditions on joke reactions, collapsed over time. In sum, these analyses centered on testing the simple mediational effects of the three-cell model on joke reactions.⁹

Direct effects of self-affirmation condition on joke reactions

To test whether type of self-affirmation predicted reactions to the black jokes, the first regression equation entered the contrast codes as predictor variables with ratings on the combined joke index as the dependent variable.

Results indicated that CC1, the contrast between the no self-affirmation condition and the average of the self-affirmation conditions, was a marginally significant predictor of joke ratings, $F(1, 93) = 3.33, p < .07$ ($\beta = -.19$), such that participants in the related and unrelated self-affirmation conditions rated the jokes as funnier than those in the no self-affirmation conditions. A second regression equation entered the two contrast codes as predictor variables, with scores on the nonverbal reaction index as the dependent variable. CC1 was a significant predictor of nonverbal reactions, $F(1, 88) = 4.69, p < .03$ ($\beta = -.19$), such that related and unrelated self-affirmation participants reacted more favorably to the jokes than no self-affirmation participants. These regression analyses are consistent with the results of the ANOVA's presented previously, and indicate that type of self-affirmation directly influenced reactions to the black jokes, such that no self-affirmation participants inhibited prejudiced responses but related and unrelated self-affirmation participants did not.

Direct effects of self-affirmation condition on affect and confidence

To test whether type of self-affirmation predicted negself, the first regression equation entered the two contrast codes as predictor variables with negself as the dependent variable. Results showed that the effect of CC1 on negself ratings was significant, $F(1, 93) = 13.49, p < .001$ ($\beta = .35$). This finding is consistent with previous analyses that showed levels of negself were higher in the no self-affirmation conditions than they were in the related and unrelated self-affirmation conditions. To test whether type of self-affirmation predicted positive

affect, a second regression equation entered the contrast codes as predictor variables with positive affect as the dependent variable. Results showed that the effect of CC1 on positive affect ratings was significant, $F(1, 93) = 9.06$, $p < .003$ ($\beta = -.26$). This finding is consistent with previous analyses that showed levels of positive affect were lower in the no self-affirmation conditions than they were in the related and unrelated self-affirmation conditions. To test whether type of self-affirmation predicted general confidence, a third regression equation entered the contrast codes as predictor variables with general confidence as the dependent variable. Results revealed that the effect of self-affirmation type on general confidence ratings was significant, $F(1, 93) = 7.68$, $p < .007$ ($\beta = .50$), consistent with previous results that found no self-affirmation participants were more confident than participants who self-affirmed. Type of self-affirmation did not significantly predict depressed affect, negative other-directed affect, or prejudice-relevant confidence.

Mediational effects of negself, positive affect, and confidence

To test whether negself mediated the effects of self-affirmation condition on joke reactions, the contrast codes and negself were entered into the regression model at the same step, as predictor variables, with joke ratings and nonverbal reactions as the dependent variables (in two separate equations). Results showed that when negself was controlled in the model, the effects of self-affirmation condition on joke ratings and nonverbal reactions were still significant. Further, negself did not significantly predict joke ratings or nonverbal reactions.

The same nonsignificant results were found for positive affect and general confidence. Overall, these results indicate that none of the affect indices and neither of the confidence items directly mediated the effects of self-affirmation condition on joke reactions.

Discussion

This study examined the affective and behavioral consequences of prejudice-relevant discrepancies and self-affirmations for low-prejudice participants. Discrepancies were activated by having participants write about why they believed it was important to respond to black people in a nonprejudiced manner, then recall two personal examples of times they failed to live up to their nonprejudiced standards. After the discrepancy manipulation, participants self-affirmed on either a related or unrelated self dimension, or did not engage in an explicit self-affirmation. The influences of the various self-affirmations on participants' affect and confidence, as well as on their self-reported and nonverbal reactions to racial jokes, were examined.

Results indicated that both related and unrelated self-affirmations alleviated the guilt associated with prejudice-relevant discrepancies and boosted positive affect. Although participants who self-affirmed were feeling better than participants who did not self-affirm, no self-affirmation participants reported higher levels of general confidence than participants who engaged in related self-affirmations. In terms of behavioral outcomes, both types of self-affirmations

disrupted the prejudice reduction process. Only participants who did not self-affirm successfully inhibited prejudiced responses to racial jokes.

Affective Consequences of Discrepancy and Self-affirmation

Alleviation of guilt

One of the central questions in this study was whether related and unrelated self-affirmations would relieve the guilt associated with prejudice-relevant discrepancies for low-prejudice people. The pattern of results provided clear evidence that this was the case.

Participants who experienced prejudice-relevant discrepancies without an explicit opportunity to self-affirm experienced high levels of negative self-directed affect (i.e., guilt) relative to participants in a no-discrepancy control group. This finding replicates previous work that has examined low-prejudice participants' affective reactions to prejudice-related discrepancies (Devine et al., 1991; Monteith et al., 1993; Monteith, 1993; Zuwerink et al., 1996; Vance et al., 1997). In contrast, participants who self-affirmed on either a related or unrelated self dimension after discrepancy activation experienced decreased levels of guilt relative to participants who did not self-affirm. Although both self-affirmations alleviated discrepancy-related guilt, they did not completely relieve guilt back to its lowest levels, as found in the control condition. This finding is somewhat inconsistent with Vance et al. (1997), who found that only unrelated self-affirmations alleviated discrepancy-related guilt. In the present study, both related and unrelated self-affirmations relieved guilt at comparable levels.

Two differences in the methodologies of these studies may help to explain this divergent finding. Vance et al. (1997) activated prejudice-relevant discrepancies by asking participants to imagine how they should and actually would respond to homosexuals in hypothetical contact situations. The present study asked participants to recall two actual instances in which they had responded negatively to black people. The absolute levels of guilt associated with discrepancies in this study were higher than in Vance et al.'s (1997) study, allowing a greater range for testing whether self-affirmations alleviated guilt. Also, the difference in the target groups across the two studies may have influenced how effectively participants were able to utilize the related self-affirmation opportunities. Participants in the present study were reflecting on their reactions to black people, a familiar minority group with whom participants have had previous opportunities for contact. Participants in Vance et al. (1997) were focused on their reactions to homosexuals, a less familiar minority group for many participants. That is, participants in these studies generally have had fewer opportunities for contact with homosexuals than with black people, and because of the "invisibility" of sexual orientation, may be less aware of instances when they interacted with homosexuals. As a result, participants in the present study may have been better able to generate examples of times they responded without prejudice than were participants in Vance et al.'s (1997) study, making the related self-affirmation more effective for the present participants.

Positive affect and confidence

In addition to testing whether self-affirmations alleviated discrepancy-related guilt, this study investigated the effects of self-affirmations on positive affect. Participants who did not self-affirm experienced decreased levels of positive affect relative to participants in a no-discrepancy control group. As was the case for negself, this finding replicates previous research examining affective reactions to prejudice-related discrepancies (Devine et al., 1991; Monteith et al., 1993; Monteith, 1993; Zuwerink et al., 1996; Vance et al., 1997). Participants who self-affirmed on either a related or unrelated self dimension after discrepancy activation experienced increased levels of positive affect relative to participants who did not self-affirm. Although both self-affirmations boosted positive affect, they did not completely restore positive affect back to its highest levels, as found in the control condition. This finding is consistent with Vance et al. (1997), who found that both related and unrelated self-affirmations increased positive affect following prejudice-related discrepancies.

Regarding the effects of self-affirmation on confidence, it was predicted based on Vance et al.'s (1997) findings that both the related and unrelated self-affirmations would increase levels of prejudice-relevant confidence over levels in the no affirmation condition. This finding was not supported. Levels of prejudice-relevant confidence did not differ significantly across affirmation conditions. A ceiling effect for this variable restricted the range of variability, possibly accounting for the nonsignificant finding in the present study. Again, the

difference in target groups across the two studies may also help explain the failure to replicate Vance et al. (1997). Participants in the present study uniformly reported high levels of confidence in their abilities to respond to black people in a nonprejudiced manner. In contrast, some of Vance et al.'s (1997) participants reported relatively low levels of confidence in their abilities to respond to homosexuals without prejudice. The greater range of responses in Vance et al.'s (1997) study may have permitted a more sensitive test of the effects of affirmation on prejudice-relevant confidence.

Interestingly, type of self-affirmation significantly affected ratings of the general confidence item, such that participants who did not self-affirm reported higher levels of general confidence than participants who self-affirmed on a related self dimension. This finding seems incongruous with results for the affect indices. Why would participants in the no self-affirmation condition, who were experiencing the most guilt and the least positive affect, report the highest levels of general confidence? Or, conversely, why would participants in the related self-affirmation condition, who were experiencing lower levels of guilt and more positive affect, report the lowest levels of general confidence? Unfortunately, it did not make sense to measure confidence in the context of the control condition, where discrepancies were not activated. As a result, there is no baseline measure of general confidence to establish whether no self-affirmation participants inflated confidence ratings, or whether related self-affirmation participants suffered a decrease in confidence.

Nonetheless, one might speculate that in the absence of an explicit self-affirmation opportunity, participants in the no self-affirmation condition attempted to use the confidence item to bolster their self-images, thereby artificially inflating confidence ratings. Alternatively, perhaps this finding belies a drawback associated with related self-affirmations. Thibodeau and Aronson (1992) cautioned that related self-affirmations might heighten discrepancy-related distress by serving as reminders of the discrepancies. The net result of the related self-affirmation may have been to shake participants' confidence even as it alleviated guilt and boosted positive affect associated with the initial discrepancy. Future research is needed to provide a clearer picture of how various self-affirmations influence general and discrepancy-related confidence, and how confidence in turn affects prejudice-relevant behavior. Such research should include multiple, reliable measures of confidence to expand on the single exploratory items used in the present study.

Behavioral Consequences of Discrepancy and Self-affirmation

Inhibition of future prejudiced responses

A second important question was whether related and unrelated self-affirmations, by eliminating guilt, the motivating force behind self-regulation efforts, would disrupt the prejudice reduction process. Results suggested that self-affirmations did indeed interfere with the inhibition of prejudiced responses.

Consistent with Monteith's (1993) findings, it was predicted that participants who experienced prejudice-relevant discrepancies without an

opportunity to self-affirm would rate jokes at the expense of black people least favorably. Evidence supporting this prediction was obtained. Participants in the no self-affirmation conditions, who experienced the highest levels of guilt, inhibited ratings of black jokes relative to participants in the no-discrepancy control condition and participants in the related self-affirmation conditions. In contrast, participants in the related and unrelated self-affirmation conditions did not inhibit joke ratings relative to the no-discrepancy control group. Results for participants' nonverbal reactions to the black jokes supported this pattern of findings, and lend credibility to the self-report measures of joke favorability.

In sum, with regard to both self-reports and nonverbal reactions, participants who did not self-affirm after prejudice-relevant discrepancies reacted most negatively to the black jokes, whereas participants who engaged in related self-affirmations reacted most favorably. These findings suggest that the only participants who engaged in prejudice reduction efforts following prejudice-relevant discrepancies were those who did not self-affirm. Related and unrelated self-affirmations disrupted the prejudice reduction process.

Facilitation of volunteer behavior

In addition to influencing the inhibition of personal prejudiced responses, self-affirmations were expected to affect the willingness of participants to volunteer to make phone calls to promote diversity efforts. Participants in the no affirmation conditions, who were feeling the most guilty about prejudice-related discrepancies, were predicted to volunteer to make the most phone calls and

were expected to volunteer at higher rates than participants in the affirmation conditions. These predictions were not supported by the data. No significant effects of type of self-affirmation or time of assessment were found for the volunteer measure. One possible reason for the nonsignificant effects is that the volunteer variable displayed a floor effect--only 17% of participants volunteered at all, and those who did volunteered to make relatively few phone calls. The narrow range of responses constrained testing the effects of the self-affirmation conditions.

Beyond this methodological issue, it is possible that the phone call measure lacked construct validity, and was not a valid measure of prejudiced behavior. For example, failure to volunteer may have reflected that a participant was shy, busy, not a confident speaker, etc., and may not have represented a "prejudiced" response at all. This problem raises an issue that regularly confronts prejudice researchers, and that is determining what behaviors best reflect "nonprejudiced" behavior. Various studies have operationalized nonprejudiced behavior differently: donating money to a black panhandler (Dutton & Lake, 1973); volunteering time to an interracial "brotherhood" campaign (Dutton & Lennox, 1974); inhibiting positive responses to racial jokes (Monteith, 1993); failing to endorse racial stereotypes (Plant & Devine, in press). Determining which behaviors best reflect prejudice-relevant attitudes is a complex matter.

Research on the correlations between attitudes and behavior has shown that general attitudes, e.g., attitudes toward black people, are typically weak

predictors of specific, single behaviors, e.g., volunteering to make phone calls to promote diversity efforts (Ajzen & Fishbein, 1977; Fishbein & Ajzen, 1975). Perhaps the joke rating task used in the present study was a more sensitive measure of prejudice-relevant behavior than number of phone calls because participants were pre-selected based, in part, on their attitudes toward finding humor in racial jokes--an attitude compatible with the specific behavior measured. The point to be gleaned from this discussion is that future research should strive to measure reliable multiple-act criterion (prejudice-relevant behaviors) that are compatible (in terms of generality versus specificity) with attitudes hypothesized to predict behavior. Number of phone calls, a specific behavior, may simply have been incompatible with general attitudes toward blacks, and therefore not a valid or sensitive measure of prejudice-relevant behavior.

The effects of time on behavioral inhibition

In addition to examining whether self-affirmation conditions influenced behavioral inhibition, this study tested how the passage of time affected inhibition of prejudiced responses. Monteith (1993) speculated that "if the discrepancy experience is strong enough to engage the self-regulatory cycle fully, inhibition may result even after a protracted period of time" (p. 482). Results of this study generally support Monteith's prediction. There were no significant time effects for the combined joke index or the nonverbal reaction index, suggesting that participants in the no self-affirmation conditions inhibited positive joke reactions both immediately and two days after the prejudice-related discrepancy

experience. In terms of combating prejudice, this finding is exciting because it suggests that low-prejudice people can learn from discrepancy experiences to control their future prejudice-relevant behavior, and that the benefits of this learning process extend over time.

On a less optimistic note, participants in the related and unrelated self-affirmation conditions failed to inhibit positive joke reactions both immediately and over time. This finding suggests that, despite experiencing meaningful prejudice-relevant discrepancies and discrepancy-related guilt, participants who self-affirmed did not engage in the prejudice reduction process. All participants in the present study, except those in the no-discrepancy control condition, presumably experienced high levels of guilt immediately after discrepancy activation. Monteith's (1993) analysis suggested that the experience of discrepancy-related guilt would set the prejudice reduction process into motion. Yet, when guilt was alleviated through self-affirmation, the process was disrupted. To fully understand these findings, we must attempt to ascertain the mechanism by which self-affirmations interfered with prejudice reduction efforts.

Mediation of Prejudice-relevant Behavior

Mediational analyses were conducted to investigate whether guilt, positive affect or confidence explained the effects of self-affirmations on prejudice-relevant behavior. It was hypothesized that guilt would directly mediate the effects of self-affirmation condition on joke reactions, consistent with Monteith's (1993) finding that guilt was necessary to initiate self-regulatory efforts following prejudice-

relevant discrepancies. It was also proposed that positive affect and confidence would mediate the inhibition of prejudiced responses by encouraging renewed efforts at behaving in a nonprejudiced manner (Carver et al., 1979; Carver & Scheier, 1990). Results showed that neither guilt, positive affect nor confidence were the mechanisms by which self-affirmations directly influenced joke reactions.

Although guilt was not a direct mediator of prejudice-relevant behavior, increased levels of guilt, experienced by no self-affirmation participants, were associated with the most nonprejudiced reactions to the racial jokes. Thus, it appears that guilt was a necessary component of the prejudice reduction process. But if guilt does not lead directly to nonprejudiced behavior, what is its role in the process? Previous research provides an answer to this question. Monteith (1993) demonstrated that guilt directly motivated low-prejudice people to self-regulate in the wake of prejudice-relevant discrepancies. Cognitive self-regulation efforts were reflected by self- and discrepancy-focused thoughts, attention to discrepancy-relevant information, and slowing of future prejudice-relevant responses. Thus, it may be the case that guilt is a necessary motivator of cognitive self-regulation, and that self-regulation directly mediates inhibition of prejudiced responses.

The present study did not explicitly measure self-regulation processes, and so cannot empirically address this proposition. However, research is currently underway to test the hypothesis that self-affirmations disrupt prejudice reduction by alleviating guilt and thereby short-circuiting the cognitive self-regulation

activities necessary to facilitate inhibition of prejudiced responses. This research is examining the effects of related and unrelated self-affirmations on cognitive self-regulatory processes, as measured in Monteith's (1993) study. Based on the findings from the present study, it is predicted that both types of self-affirmations will disrupt self-regulation following prejudice-related discrepancies, relative to a no self-affirmation condition, with related self-affirmations disrupting the process more completely than unrelated self-affirmations.

Alternative Explanations

Inhibited joke ratings as self-affirmation

It has been argued that participants who did not self-affirm following prejudice-relevant discrepancies inhibited their subsequent reactions to racial jokes because the increased guilt they experienced motivated them to self-regulate and engage in prejudice reduction efforts. An alternative explanation is that, in the absence of an explicit self-affirmation opportunity, these participants seized upon the joke rating task as an opportunity to restore their nonprejudiced identities. Essentially, no self-affirmation participants may have used the joke rating task as a kind of related self-affirmation. Inhibited joke ratings may therefore reflect efforts to self-affirm rather than providing evidence of self-regulation and prejudice reduction.

In a conceptually relevant study, Dutton and Lake (1973) gave low-prejudice participants false physiological feedback that led participants to believe they might be prejudiced toward blacks. After leaving the laboratory, one group of

participants was panhandled by a black confederate and another group was panhandled by a white confederate. Results showed that when the panhandler was black, participants donated an average of \$.47, with 85% of participants donating some amount. In contrast, when the panhandler was white, an average of only \$.28 was donated, with 50% of participants donating some amount. Low-prejudice participants who did not experience a prejudice-relevant discrepancy donated an average of \$.17 to a black panhandler and \$.28 to a white panhandler. Dutton and Lake interpreted these findings to mean that participants who were panhandled by the black confederate used that opportunity to prove to themselves through their behavior that they were not prejudiced, despite the physiological feedback. Donating to the black panhandler represented an easy way of reestablishing their nonprejudiced self-images.

In the context of the present study, it is plausible that no self-affirmation participants viewed the joke rating task as a chance to redeem their nonprejudiced self-images. These participants may have reacted negatively to the racial jokes as a way of demonstrating that they were, in fact, nonprejudiced, just as Dutton and Lake's participants demonstrated their egalitarianism by donating to a black panhandler. If the joke rating task served as a sort of self-affirmation for no affirmation participants, there are potential consequences that require further investigation. First, it would be of interest to know whether the joke rating task relieved discrepancy-related guilt for no self-affirmation participants, much as the related and unrelated self-affirmations in the present study did.

Second, if the joke rating task did relieve guilt, it would be important to know whether it also interfered with future prejudice-relevant behavior, just as the self-affirmations in the present study disrupted inhibited joke ratings.

Previous research has investigated this latter issue. A study conducted by Dutton and Lennox (1974) showed that not only do low-prejudice people seek to directly restore their nonprejudiced self-images following prejudice-relevant discrepancies, but that doing so disrupts long-term prejudice reduction efforts. As in Dutton and Lake (1973), a false physiological feedback procedure was used to activate prejudice-relevant discrepancies for low-prejudice participants. Following discrepancy activation, participants were either panhandled by a black confederate, a white confederate, or not panhandled at all. Two days later, all participants were contacted and asked if they would be willing to donate time to an interracial "brotherhood" campaign. Dutton and Lennox predicted that participants who donated money to the black panhandler, thereby reestablishing their nonprejudiced identities, would commit themselves to less effort on the campaign than participants who were not panhandled by a black person. Results supported this prediction. Participants in the black panhandler condition committed themselves to significantly less effort than participants in the white panhandler and no panhandler conditions (which were not significantly different from one another). Dutton and Lennox concluded that by donating to the black panhandler, participants were able to reestablish their nonprejudiced self-images,

thus "freeing" themselves from the need to engage in more long-term, effortful prejudice reduction efforts.

Dutton and Lennox (1974) used these findings to argue that "token" nonprejudiced behaviors may reassure low-prejudice people that they are, in fact, nonprejudiced, thereby relieving them of the need to undertake more significant prejudice reduction efforts. In the present study, this argument may help explain why so few participants volunteered to make phone calls following the joke rating task. With their self-images restored, either by way of explicit self-affirmation or use of the joke rating task, participants may not have been motivated to commit time to supporting diversity efforts on campus.

Results for no self-affirmation participants in the delayed assessment condition seem inconsistent with the joke-ratings-as-self-affirmation hypothesis. These participants, who experienced heightened levels of guilt and decreased positive affect following discrepancy activation, inhibited prejudiced responses to the racial jokes two days after the discrepancy experience. Over the two day period, it seems possible that no self-affirmation participants might have found opportunities outside the lab to restore their nonprejudiced self-images. As Dutton and Lake (1973) demonstrated, this can be done as easily as giving some change to a black panhandler. Yet despite the opportunities participants may have encountered for restoring their nonprejudiced identities after leaving the lab, they still inhibited positive reactions to the racial jokes two days later. If this line of reasoning is correct, it suggests that inhibited joke ratings in the delayed

condition reflected more than attempts to self-affirm. Instead, discrepancy-related guilt may have triggered self-regulation efforts that promoted nonprejudiced responses over time. It will be important for future research to empirically test whether responses to the present joke rating task were a strategic form of self-affirmation for participants with no other explicit self-affirmation opportunity, or an indicator of more permanent prejudice-relevant behavioral regulation. To this end, research is needed that will extend the present methodology by a) measuring the effects of self-affirmation on self-regulation, as described previously, and b) measuring affective and behavioral consequences for no self-affirmation participants after they complete the joke rating task.

Self-affirmations as distractions

The present findings have been used to advance the argument that related and unrelated self-affirmations disrupted prejudice reduction efforts by relieving guilt. An alternative explanation is that the self-affirmations distracted participants from thinking about the discrepancy experience, and that memory effects were responsible for elevated joke ratings, not alleviated guilt. That is, participants who self-affirmed may not have remembered the recent discrepancy experience as well as those who did not self-affirm. Future research could test this competing hypothesis by adding a control condition to the current design in which participants would engage in an activity, following discrepancy activation, that was distracting but not self-affirming. For example, participants might write about four instances in which a friend exemplified an important quality. Such an activity

should be as distracting as the present self-affirmation opportunities, but would not be expected to relieve guilt. If results showed that participants in the distraction condition, who would presumably still be feeling guilty, failed to inhibit joke ratings, this would support the hypothesis that self-affirmations disrupt prejudice-relevant behavior via distraction rather than through alleviating guilt.

Although it is important to test this competing explanation, the distraction hypothesis is not inconsistent with the present analysis. It may be the case that self-affirmations disrupt prejudice reduction efforts by alleviating discrepancy-related guilt and distracting people from reflecting on the discrepancy experience. The net effect of guilt alleviation and distraction might be to interfere with self-regulation processes. A complete model of the variables associated with self-affirmation and prejudice reduction is likely to include cognitive components (like distraction and self-regulation) as well as the affective and behavioral components which were the focus of the present study.

Theoretical Implications and Conclusions

Prejudice reduction

This study revealed a technique that was effective at promoting prejudice reduction efforts. Asking low-prejudice participants to publicly write about why they believed it was important to treat black people fairly, then to recall times they personally failed to live up to this standard, created meaningful prejudice-related discrepancies. In the absence of self-affirmation, these discrepancies aroused feelings of guilt that promoted inhibition of prejudiced responses even two days

later. Prejudice reduction “practitioners” are encouraged to make use of this finding as they develop strategies for helping individuals and organizations reduce discrimination and prejudice. Encouraging low-prejudice people to think deeply about their nonprejudiced standards, then to reflect on situations in which they violated these standards, may be a useful technique for promoting prejudice reduction. Future research efforts will be aimed at testing whether this prejudice-related discrepancy manipulation will promote inhibition of prejudiced responses in applied settings. For example, this paradigm could be incorporated into diversity training workshops in large organizations and the military, and subsequent prejudice-relevant behavior compared for units and individuals who did or did not participate in the training.

It is important to note that this technique was only effective at reducing prejudiced responses when self-affirmations were not used to alleviate discrepancy-related guilt. Results showed that affirmation of related and unrelated self domains alleviated the guilt low-prejudice people typically experience when they violate their internalized, self-defining nonprejudiced standards. By reducing the motivation to inhibit prejudiced responses, both related and unrelated self-affirmations disrupted immediate and long-term prejudice reduction efforts. The results of this study highlight an important threat to the prejudice reduction process: self-affirmation. Well-intentioned people who desire to be low in prejudice may fail to learn to control prejudiced responses if they choose to relieve guilt by self-affirming instead of doing the important work of

self-regulation. The present findings suggest the need to caution low-prejudice individuals about the dangers of self-affirmation. It appears to be important for people to experience the full magnitude of guilt and discomfort associated with prejudice-relevant discrepancies so that the self-regulatory process may be fully engaged. Diversity training programs that try to make people feel good about, or distract people from thinking about, previous prejudice-relevant failures may inadvertently short-circuit prejudice reduction efforts.

Self-affirmation theory

Although this study showed that the related and unrelated self-affirmation opportunities explicitly provided to participants interfered with prejudice reduction efforts, it is interesting to speculate about whether people spontaneously self-affirm following discrepancies in the "real world". The fact that no self-affirmation participants in the present study still inhibited prejudiced responses two days post-discrepancy suggests that they might not have spontaneously self-affirmed. If they had spontaneously self-affirmed shortly after leaving the lab, thereby relieving guilt and restoring positive affect, it seems the self-regulatory process would not have been fully activated for these participants. And yet, no self-affirmation participants were still motivated to control prejudiced responses two days later. Why didn't no self-affirmation participants spontaneously self-affirm?

To understand whether self-affirmations pose a practical threat to real-world prejudice reduction efforts, we must attempt to ascertain the conditions under which people spontaneously self-affirm. Perhaps there are individual

difference variables which make certain kinds of people more likely to use self-affirmation. For example, Steele, Spencer and Lynch (1993) found that people with high self-esteem were more likely to self-affirm than people with low self-esteem because they have many favorable self-aspects to draw upon. People with low self-esteem, in contrast, were more likely to directly address the discrepancy because they lacked the esteem resources necessary to self-affirm. If this is the case, self-affirmations may only threaten prejudice reduction efforts for low-prejudice people with high self-esteem. Low-prejudice people with low self-esteem may be naturally inclined to self-regulate to learn to avoid future discrepancies.

There may be situational variables as well that influence the likelihood that self-affirmation would be employed. For example, perhaps people try to directly resolve discrepancies unless the threat posed by a particular discrepancy is intolerable. If discrepancy-related guilt exceeds a tolerable threshold, perhaps people would tend to self-affirm as an easy, immediate way of relieving guilt. In the present study, perhaps private recollection of failures to live up to nonprejudiced standards aroused guilt, but was not so intolerable to participants that they felt the need to spontaneously self-affirm. Instead, in the absence of an explicit self-affirmation opportunity, participants chose to inhibit prejudiced responses as a way of directly addressing their prejudice-relevant discrepancies. However, public failure to live up to nonprejudiced standards may pose a more intolerable threat to low-prejudice people, particularly those who are sensitive to

external evaluative sources (Plant & Devine, in press). When violations of nonprejudiced standards are made public, low-prejudice people might be more likely to self-affirm as an immediate way of relieving extreme and intolerable distress. It will be fruitful for future research to identify the conditions under which people spontaneously self-affirm, for it is only under these conditions that self-affirmations threaten to disrupt behavioral regulation.

The findings of the present study are provocative beyond the insights they provide about the affective and behavioral consequences of self-affirmation in the context of prejudice reduction. Although this study was primarily interested in understanding how self-affirmations influence the prejudice reduction process, the results also bear on an existing debate in the self-affirmation literature: whether related or unrelated self-affirmations are most effective at reducing discrepancy-related distress. Results suggest that both related and unrelated self-affirmations comparably alleviated discrepancy-related guilt and discomfort, and boosted positive affect in the wake of a prejudice-relevant discrepancy experience. Thus, in terms of countering discrepancy-related affect, both types of affirmations were found to be equally effective.

The results of this study also highlight the need for more specific criteria to define what is meant by self-affirmation "effectiveness". Do self-affirmations that "work" simply relieve psychological discomfort and increase positive affect, as did the self-affirmations in this study? It is suggested that this view is incomplete. Rather, in determining whether a self-affirmation "works" for a person, one must

consider the tradeoff between short-term psychological relief, which self-affirmations provided in this study, and long-term behavioral adjustment, which self-affirmations threatened. Would we consider a self-affirmation "effective" if it made a person with a prejudice-relevant discrepancy feel better about his or her behavior in the short-term, but left him or her vulnerable to continuing to respond with prejudice in the long-term? The current self-affirmation literature offers scant empirical evidence regarding the implications of self-affirmation for adaptive behavior. However, as this study has demonstrated, behavioral consequences of self-affirmation are perhaps even more important in the long run than affective consequences, at least when the behaviors in question are meaningful to individuals and society at large.

References

- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*, 888-918.
- Aronson, E. (1968). The theory of cognitive dissonance: A current perspective. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol 4, pp. 1-34). San Diego, CA: Academic Press.
- Aronson, J., Blanton, H., & Cooper, J. (1995). From dissonance to disidentification: Selectivity in the self-affirmation process. *Journal of Personality and Social Psychology*, *68*, 986-996.
- 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.
- Blanton, H., Cooper, J., Skurnik, I., & Aronson, J. (1997). When bad things happen to good feedback: Exacerbating the need for self-justification with self-affirmations. *Personality and Social Psychology Bulletin*, *23*, 684-692.
- Brigham, J. C. (1993). College students' racial attitudes. *Journal of Applied Social Psychology*, *23*, 1933-1967.
- Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*, 19-35.
- Carver, C. S., Blaney, P. H., & Scheier, M. F. (1979). Reassertion and giving up: The interactive role of self-directed attention and outcome expectancy. *Journal of Personality and Social Psychology*, *37*, 1859-1870.
- 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., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology*, *60*, 817-830.
- Devine, P. G., & Monteith, M. J. (1993). The role of discrepancy associated affect in prejudice reduction. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in intergroup perception*, (pp. 317-344). New York: Academic Press.

- Devine, P. G. (1995). Prejudice and outgroup perception. In A. Tesser (Ed.), *Advanced social psychology*, (pp. 467-524). New York: McGraw-Hill.
- Devine, P. G., Evett, S. R., & Vasquez-Suson, K. A. (1996). Exploring the interpersonal dynamics of intergroup contact. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: The interpersonal context* (Vol 3). New York: Guilford.
- Dickerson, C., Thibodeau, R., Aronson, E., & Miller, D. (1992). Using cognitive dissonance to encourage water conservation. *Journal of Applied Social Psychology*, 22, 841-854.
- Dutton, D. G., & Lake, R. A. (1973). Threat of own prejudice and reverse discrimination in interracial situations. *Journal of Personality and Social Psychology*, 28, 94-100.
- Dutton, D. G., & Lennox, V. L. (1974). Effect of prior "token" compliance on subsequent interracial behavior. *Journal of Personality and Social Psychology*, 29, 65-71.
- Festinger, L. (1957). An introduction to the theory of dissonance. *A theory of cognitive dissonance* (pp. 1-31). Stanford: Stanford University Press.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley, 1975.
- Freud, S. (1928). Humour. *International Journal of Psychoanalysis*, 9, 1-6.
- Fried, C. B., & Aronson, E. (1995) Hypocrisy, misattribution, and dissonance reduction. *Personality and Social Psychology Bulletin*, 21, 925-933.
- Fried, C. B. (in press). Hypocrisy and identification with transgressions: A case of undetected dissonance. *Basic and Applied Social Psychology*.
- Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, 43, 349-358.
- Guilford, J. P. (1954). *Psychometric Methods*. New York: McGraw-Hill.
- Hardyck [Piliavin], J. A. (1966). Consistency, relevance, and resistance to change. *Journal of Experimental Psychology*, 2, 27-41.

- Hardyck [Piliavin], J. A., & Kardush, M. (1968). A modest modish model for dissonance reduction. In Abelson, R. P., Aronson, E., McGuire, W. J., Newcomb, T. M., Rosenberg, M. J., & Tannenbaum, P. H., (Eds.), *Theories of Cognitive Consistency: A Sourcebook*. Chicago: Rand McNally.
- LaFrance, M., & Woodzicka, J. A. (in press). No laughing matter: Women's Verbal and nonverbal reactions to sexist humor. To be included in: J. Swim and C. Stangor (Eds.), *Targets of Prejudice*.
- McGhee, P. E. (1988). The contribution of humor to children's social development. *Journal of Children in Contemporary Society*, 20, 119-134.
- 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., 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.
- Neitz, M. J. (1980). Humor, hierarchy, and the changing status of women. *Psychiatry*, 43, 211-223.
- Newman, R. J. (March 10, 1997). Did we say zero tolerance? *U. S. News & World Report*, 33-34.
- Plant, E. A., & Devine, P. G. (in press). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*.
- Reibstein, L. (November 25, 1996). Managing diversity. *Newsweek*, 50.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Sherman, S. J., & Gorkin, L. (1980). Attitude bolstering when behavior is inconsistent with central attitudes. *Journal of Experimental Social Psychology*, 16, 388-403.
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. *Advances in Experimental Social Psychology*, 21, 261-300.
- Steele, C. M., Spencer, S. J., & Lynch, M. (1993). Self-image resilience and dissonance: The role of affirmational resources. *Journal of Personality and Social Psychology*, 64, 885-896.

- Stone, J., Aronson, E., Crain, A. L., Winslow, M. P., & Fried, C. B. (1994). Inducing hypocrisy as a means of encouraging young adults to use condoms. *Personality and Social Psychology Bulletin*, 20, 116-128.
- Stone, J., Wiegand, A., Cooper, J., & Aronson, E. (1997). When exemplification fails: Hypocrisy and the motive for self-integrity. *Journal of Personality and Social Psychology*, 72, 54-65.
- Tabachnik, B. G., & Fidell, L. S. (1996) *Using Multivariate Statistics* (3rd Edition). New York: Harper Collins.
- Tangney, J. P. (1995). Recent advances in the empirical study of shame and guilt. *American Behavioral Scientist*, 38, 1132-1145.
- Thibodeau, R., & Aronson, E. (1992). Taking a closer look: Reasserting the role of the self-concept in dissonance theory. *Personality and Social Psychology Bulletin*, 18, 591-602.
- Vance, K. M., Devine, P. G., & Barron K. E. (1997, May). *The Effects of Self-affirmation on the Prejudice Reduction Process*. Paper presented at the Sixty-ninth Annual Meeting of the Midwestern Psychological Association Conference, Chicago, IL..
- Vistica, G. (November 25, 1996). Rape in the ranks. *Newsweek*, 28-31.
- Zuwerink, J. R., Devine, P. G., Monteith, M. J., & Cook, D. A. (1996). Prejudice toward Blacks: With and without compunction? *Basic and Applied Social Psychology*, 18 (2), 131-150.

Footnotes

1. Data from 43 additional participants were excluded from analyses for the following reasons. A prejudice-relevant discrepancy was not fully activated for 23 participants. Six participants completed the behavioral measures more than two days after discrepancy activation. Seven participants generated incomplete materials. Three participants suspected that the behavioral measures were related to the discrepancy manipulation. And four of the control group participants evidenced unusually high baseline levels of negative affect.
2. Two of the participants chose to write about the quality of being nonprejudiced, and their data were included in the related affirmation condition for all analyses.
3. The original item from Rosenberg's (1965) self-esteem scale read: "I am able to do things as well as most other people."
4. Only one participant had heard about the study from a friend, and his data were excluded from all analyses.
5. The two-way ANOVA performed on the prejudice-relevant confidence item revealed an unexpected main effect for time, $F(1, 90) = 4.57, p < .04$, such that participants in the immediate time conditions reported slightly lower prejudice-relevant confidence ($M = 13.48$) than participants in the delayed time conditions ($M = 14.04$). Because prejudice-relevant confidence was assessed before the time manipulation, this effect was attributed to random variance between groups. Nonetheless, the main effect for time made further interpretation of this variable, and potentially the other dependent variables, problematic. Therefore, the effect of prejudice-relevant confidence was covaried out of all further analyses.
6. Because negself and discomfort were highly correlated (Pearson's $r = .78, p < .001$), the ANOVA for negself was also run with discomfort as a covariate. Despite the significance of the covariate ($p < .001$), the main effect of self-affirmation type remained significant, $F(2, 88) = 3.22, p < .05$, suggesting that the effects of self-affirmation type on negself are distinct from those on global discomfort.
7. An interesting issue that surfaced when joke ratings were analyzed individually was that affirmation conditions influenced each of the four black jokes differently. Joke 10, a joke about ebonics, showed significant condition effects that paralleled findings for the combined joke index. Joke 1, a joke about Haitians, revealed a significant condition by time interaction, such that participants in the unrelated self-affirmation conditions rated the joke more favorably in the delayed assessment condition than they did when joke ratings were assessed immediately. Finally, Joke 5 (a joke about Stevie Wonder) and Joke 15 (a joke

about Father's Day) showed no significant effects of type of self-affirmation or time of assessment. When combined into an index, ratings of the four jokes on average converged to support the predicted relationships between self-affirmations and prejudice-relevant behavior. However, it is important to consider why the various jokes yielded different patterns of responses individually. One possibility is that the jokes reflected different dimensions of the complex phenomenon known as "prejudice". Jokes 5 and 15 both blatantly played on traditional stereotypes of black people. Joke 1 incorporated the issue of immigration, and may have conjured up stereotypes of other minority groups (e.g., Hispanics) in addition to black people. And Joke 10 raised political issues surrounding the ebonics debate. It is noteworthy that low-prejudice people reacted differently to prejudice-relevant stimuli depending on what aspects of prejudice were being tapped.

8. Although nonsignificant, these data did show an interesting trend, whereby participants in the related self-affirmation conditions volunteered less often (9%) than participants in the no self-affirmation (16%) and unrelated self-affirmation (25%) conditions.

9. Mediation analyses were also conducted with a four-cell model, which included the three self-affirmation conditions collapsed over time and the no-discrepancy control condition. Three orthogonal contrasts were constructed to represent this four-cell model in the regression equations. In addition to testing the simple mediation effects of affect/confidence on joke reactions, the mediational effects of affect/confidence in interaction with self-affirmation condition were tested for both the three- and four-cell models. Results of all these mediational analyses were consistent with those reported for the simple mediation effects of the three-cell model, as described in the "Results" section.

Table 1

Affect and Confidence Ratings in Three Pre-testing Conditions: Baseline (no discrepancy), Standards-only (essay) and Discrepancy (essay and examples)

<u>Affect index</u>	<u>Baseline</u>	<u>Standards-only</u>	<u>Discrepancy</u>
Negself	1.81 _a	1.75 _a	3.47 _b
Discomfort	2.04 _a	2.39 _{ab}	3.37 _b
Positive affect	4.73 _a	5.16 _a	2.83 _b
Negoother	2.22	3.86	2.71
Depressed	1.31	1.76	2.00
Prej-related confidence	----	13.71	14.19
General confidence	----	12.64	13.31

Affect ratings were made on a 1 ("does not apply at all") to 7 ("applies very much") scale.

Confidence ratings were made on a 1 ("not at all confident") to 15 ("extremely confident") scale.

Means in each row not sharing common subscripts differ significantly from each other (Bonferroni post hoc test, $p < .05$).

Table 2

Pre-testing Data for the Four Black Jokes: Should Ratings, Would Ratings and Discrepancy Scores

<u>Joke</u>	<u>Would</u>	<u>Should</u>	<u>Discrepancy</u>
Joke 1 (Haitians)	3.19	2.00	1.19
Joke 5 (Stevie Wonder)	3.19	1.78	1.41
Joke 10 (Ebonics)	4.19	2.70	1.49
Joke 15 (Father's Day)	3.11	1.89	1.22

Joke ratings were made on a 1 ("not at all funny") to 8 ("extremely funny") scale. Would ratings reflected how funny participants said they actually would find the jokes. Should ratings reflected how funny participants said they should find the jokes, according to their personal nonprejudiced standards. Discrepancy scores reflected the difference between would and should ratings. Positive discrepancy scores suggested actual joke ratings were somewhat prejudiced and could be inhibited to be more in line with nonprejudiced standards.

Table 3

Affect and Confidence Ratings in the Control (no discrepancy), No Self-affirmation, Unrelated Self-affirmation and Related Self-affirmation Conditions

Affect index	Control	No Affirm	Unrelated	Related
Negself	1.51 _a	4.15 _b	3.05 _c	3.13 _c
Discomfort	1.81 _a	3.46 _b	2.69 _c	2.92 _{bc}
Positive affect	4.97 _a	3.21 _b	4.16 _c	3.81 _c
Negother	1.54 _a	2.99 _b	3.00 _b	2.91 _b
Depressed	1.52 _a	2.37 _a	2.15 _a	2.20 _a
Prej-related confidence	----	13.94 _a	13.66 _a	13.69 _a
General confidence	----	13.13 _a	11.84 _{ab}	11.41 _b

Affect ratings were made on a 1 ("does not apply at all") to 7 ("applies very much") scale. Means in each row not sharing common subscripts differ significantly from each other (planned comparisons, $p < .04$).

Confidence ratings were made on a 1 ("not at all confident") to 15 ("extremely confident") scale. Means in each row not sharing common subscripts differ significantly from each other (Bonferroni post hoc test, $p < .05$).

Table 4

Joke Ratings, Nonverbal Reactions, and Number of Phone Calls in the Control (no discrepancy), No Self-affirmation, Unrelated Self-affirmation and Related Self-affirmation Conditions

<u>Measure</u>	<u>Control</u>	<u>No affirm</u>	<u>Unrelated</u>	<u>Related</u>
Joke Index	3.53 _a	2.77 _b	3.11 _{ab}	3.54 _a
Nonverbal index	1.15 _{ab}	.75 _a	1.19 _{ab}	1.40 _b
Number of calls	1.83 _a	1.55 _a	2.00 _a	1.08 _a

Joke ratings were made on a 1 ("not at all funny") to 8 ("extremely funny") scale.

Scores on the nonverbal index represented the number of black jokes to which participants reacted favorably, and ranged from 0 (no positive reactions) to 4 (positive reactions to all four jokes).

Number of phone calls was an open-ended measure; participants reported how many 5-10 minute phone calls they were willing to make to promote diversity efforts.

Means in each row not sharing common subscripts differ significantly from each other (planned comparisons, $p < .09$ for joke index; $p < .05$ for nonverbal index).

APPENDIX A

PARTICIPANT SELECTION MEASURES

ATB Scale

Please indicate your agreement/disagreement with each of the following statements, using the scale below. Please write your rating in the blank to the left of each statement.

1	2	3	4	5	6	7
Strongly Disagree						Strongly Agree

- ___ 1. I would rather not have blacks live in the same apartment building I live in.
- ___ 2. I get very upset when I hear a white make a prejudicial remark about blacks.
- ___ 3. Black and white people are inherently equal.
- ___ 4. I would not mind at all if a black family with about the same income and education as me moved in next door.
- ___ 5. It would not bother me if my new roommate was black.
- ___ 6. Whites should support blacks in their struggle against discrimination and segregation.
- ___ 7. If a black were put in charge of me, I would not mind taking advice and direction from him or her.
- ___ 8. I think that black people look more similar to each other than white people do.
- ___ 9. The federal government should take decisive steps to override the injustices blacks suffer at the hands of local authorities.
- ___ 10. I would probably feel somewhat self-conscious dancing with a black in a public place.
- ___ 11. Interracial marriage should be discouraged to avoid the "who-am-I?" confusion which the children feel.
- ___ 12. Black people are demanding too much too fast in their push for equal rights.

APPENDIX B

DISCREPANCY MANIPULATION

INSTRUCTIONS GIVEN TO PARTICIPANTS:

The experimenter greets the participant, asks him or her to complete a consent form, then says:

"Thanks for coming in today. My name is _____, and I'm a coordinator for the SOAR program. We're conducting a project in conjunction with Dr. Trish Devine in the psychology department. Because of her expertise in studying prejudice, SOAR has asked Dr. Devine to help us strengthen our segment on diversity and human relations. You probably attended SOAR this past summer, and you may remember that one of the goals of SOAR is to make new students at UW aware of the great diversity that exists within the student body. Recent incidents of discrimination on campus have suggested we need to improve our efforts to promote fair treatment of minority students. One way of doing this is to help incoming students understand why it's important to treat other students fairly, regardless of race, ethnicity, gender, sexual orientation, and religious beliefs.

Research has shown that young adults are more open to information about discrimination when it comes from peers as opposed to authority figures. The purpose of this project is to gather peer input for the diversity segment of the 1998 SOAR program. Your advice will be viewed as very credible to incoming students because you're a little older and more experienced than they are, yet not so different that you can't relate to them.

SOAR is designing a brochure that will feature quotes from students discussing their personal beliefs about why black, Asian American, and homosexual students should be treated fairly. These are the minority groups against which most of the recent discrimination has been targeted. You may remember getting a folder like this one when you attended SOAR, filled with brochures like these. *[Experimenter shows demonstration folder and opens a brochure to give participant the idea.]* The brochure we're creating will be added to this folder.

We would like you to help convey the message that blacks on campus should be treated fairly, if that's what you really believe. Is it? *[If participant says "yes", experimenter says "great" and continues. If participant says "no", experimenter thanks her or him, gives participant an extra credit point, and excuses participant.]* We'd like you to write a brief essay about why you personally believe it's important to treat black people fairly. Don't worry, your language and style can be informal. Please base your comments on your own personal standards for how you think blacks should be treated and why.

I'll give you about ten minutes to think about your personal standards and write your essay. It may be helpful to outline your ideas first, and you'll find scrap

paper on your desk. Please pull out the next form from your envelope--the "essay" form. On it, you'll find a standard introduction that we'd like everyone to use to begin their essays. Go ahead and write your essay on that page as well, and feel free to use extra pages if you like. When you're finished, please let me know."

When the participant finishes, the experimenter says:

"I'm sure your essay is great; it'll be very useful to SOAR. Now, please pull out the next form from your envelope--the "release" form. I need to ask you to fill out this release form to give SOAR permission to present portions of your essay in the brochure to be handed out to incoming students. Please take a minute to read and sign this form."

After the participant signs the release form, the experimenter says:

"Let's move on to the second part of this project. Although people often believe they should treat blacks fairly, they also find that sometimes their reactions are more prejudiced than they personally believe they should be. For example, you may have felt uncomfortable interacting with a black person or had stereotypic thoughts about a black person in situations when you think you shouldn't have."

Please take a few moments and briefly write about two situations in which you reacted more negatively to a black person than you thought you should, or treated a black person in a prejudiced manner. Your written examples will be used to spark small group discussions during the SOAR diversity segment. Your name will be kept anonymous during these discussions. Please pull the next form out of the envelope--the "personal examples" form--and use it to write your two examples. Remember, this will be kept completely confidential, and it's important that you respond openly and honestly."

When the participant finishes, the experimenter says:

No affirmation condition:

"Next, there is a brief questionnaire we would like you to complete. Please pull the last form out of the envelope. This questionnaire will be kept completely confidential, and it's important that you respond openly and honestly. Please read the instructions carefully." (*The experimenter gives the participant the affect questionnaire.*)

Unrelated affirmation condition:

"There is one last input we would like you to make to SOAR. We are interested in showing new students the wide range of personal qualities represented among students on campus. We'd like you to choose your most treasured quality, and briefly describe examples of how you demonstrate this quality in your life. As was the case with the other examples you wrote, these will be kept anonymous. Please pull the next form out of your envelope--the "Personal Qualities" form. Read the instructions carefully, then write your examples."

When the participant finishes the self-affirmation, the experimenter hands her the affect questionnaire and says:

"Finally, there is a brief questionnaire we would like you to complete. Please pull the last form out of the envelope. This questionnaire will be kept completely confidential, and it's important that you respond openly and honestly. Please read the instructions carefully."

Related affirmation condition:

"There is one last input we would like you to make to SOAR. We are interested in showing new students that many of their peers have been successful at treating others in an unbiased fashion. We'd like you to briefly describe examples of times you were successful at being nonprejudiced. As was the case with the other examples you wrote, these will be kept anonymous. Please pull the next form out of your envelope--the "Personal Qualities" form. Read the instructions carefully, then write your examples."

When the participant finishes the self-affirmation, the experimenter hands him the affect questionnaire and says:

"Finally, there is a brief questionnaire we would like you to complete. Please pull the last form out of the envelope. This questionnaire will be kept completely confidential, and it's important that you respond openly and honestly. Please read the instructions carefully."

When the participant finishes the affect questionnaire, the experimenter says:

Immediate follow-up condition:

"Well, that's the end of the session. Thank you very much for helping us out today. Because I'm with SOAR, I can't give extra credit points. But you'll be given two points for your time today after you complete the other study that was scheduled along with this one. *[Experimenter acts like she's consulting the*

schedule.] I see here that's in just a few minutes. The psych people asked me to remind you to go to Room 251. The experimenter for that study will meet you there in a few minutes, I guess."

Delayed follow-up condition:

"Well, that's the end of the session. Thank you very much for helping us out today. Because I'm with SOAR, I can't give extra credit points. But you'll be given two points for your time today after you complete the other study that was scheduled along with this one. *[Experimenter acts like he's consulting the schedule.]* I see here that's scheduled for (time, e.g., 3:00) on (day, e.g., Wednesday). The psych people asked me to remind you to go to Room 251 on (day) to wait for that study. The experimenter will meet you there, I guess."

ESSAY FOR DIVERSITY BROCHURE
(SOAR 1998)

"My name is _____, and I'm a
(your first and last name)

_____ at the University of Wisconsin, Madison.
(freshman, sophomore, etc.)

I'm originally from _____.
(your hometown)

I think it's important to treat blacks on campus fairly because . . ."

Student Orientation, Advising, and Registration (SOAR)

*Student Orientation Programs
Gordon Commons Offices
717 W. Johnson Street
Madison, WI 53706-1790*

**RELEASE FORM FOR DIVERSITY ESSAY
(SOAR 1998)**

I understand that I have freely chosen to write an essay that will be used as part of the materials for the segment on "Diversity and Human Relations" in the 1998 SOAR program. I understand that at least part of my essay will be published in a brochure that will be presented to incoming students at SOAR 1998. I agree to allow my essay to be included in the brochure, and to have my name associated with my essay.

Name (please print):

Signature:

Date:

**PERSONAL EXAMPLES FOR SMALL GROUP DISCUSSION
(SOAR 1998)**

Although people often believe they should treat blacks fairly, they also find that sometimes their reactions are more negative than they personally believe they should be. For example, you may have felt uncomfortable interacting with a black person or had stereotypic thoughts about a black person in situations when you think you shouldn't have.

Please take a few moments and briefly write about two situations in which you reacted more negatively to a black person than you thought you should. For example, you may want to reflect on the first time you met, shook hands with or worked with a black person. Or you might consider your reactions to blacks as portrayed in the media.

Personal Example 1:

Personal Example 2:

INSTRUCTIONS - READ VERY CAREFULLY

At this point, we would like you to reflect back on the examples you just gave for how you have reacted to blacks in the past. We are interested in how you are **feeling** about how well these actual responses to blacks, the behavior described in your examples, match your personal standards for treating blacks, as presented in your essay. Think about what words best summarize how you are feeling.

In the blanks below, please list the **three** words that best describe how you are feeling about how well your actual responses (your personal examples) match your personal standards (your essay). Please put one feeling word in each of the blanks below.

APPENDIX C

SELF-AFFIRMATION MATERIALS

**PERSONAL QUALITIES DISCUSSION MATERIAL
(SOAR 1998)**

SOAR is also interested in showing new students the wide range of personal qualities represented among students on campus. Think about your personal qualities or characteristics that you consider to be important (for example, these might be traits, values, or skills). Which of these qualities or characteristics do you like best about yourself? If there are several qualities you like about yourself, choose what you consider to be one of the most important. Please write this quality in the blank below:

List 4 examples of times when you demonstrated this quality. If you can't think of 4 examples of times when you personally demonstrated this quality, list 4 general examples of things that you think a person with this quality might do.

Please put only one thought or example in a box. Explain it in enough detail so that someone else could easily imagine your example. Don't worry about proper grammar or spelling.

**PERSONAL QUALITIES DISCUSSION MATERIAL
(SOAR 1998)**

SOAR is also interested in showing new students that many of their peers have been successful at treating others in a **NONPREJUDICED** (open-minded, unbiased, tolerant of groups) manner.

List 4 examples of times when you were successful at being **NONPREJUDICED**. If you can't think of any personal examples of times when you were successful at being nonprejudiced, list 4 general examples of things that you think a nonprejudiced person might do.

Please put only one thought or example in a box. Explain it in enough detail so that someone else could easily imagine your example. Don't worry about proper grammar or spelling.

APPENDIX D

AFFECT AND CONFIDENCE MEASURES

INSTRUCTIONS FOR NO AFFIRMATION CONDITIONS:

To provide a more complete picture of how you are feeling after writing your essay and describing your personal examples, please indicate how much each of the words below describes how you are feeling by circling a number on the scale. "1" means "does not apply at all", and "7" means "applies very much" to how you are feeling. Don't spend much time thinking about each word, just give a quick, gut-level response. It is important that you respond openly and honestly.

INSTRUCTIONS FOR UNRELATED AFFIRMATION CONDITIONS:

We are interested in how you are feeling after writing your essay, your personal examples, and describing your personal quality. Please indicate how much each of the words below describes how you are feeling by circling a number on the scale. "1" means "does not apply at all", and "7" means "applies very much" to how you are feeling. Don't spend much time thinking about each word, just give a quick, gut-level response. It is important that you respond openly and honestly.

INSTRUCTIONS FOR RELATED AFFIRMATION CONDITIONS:

We are interested in how you are feeling after writing your essay and the two sets of personal examples. Please indicate how much each of the words below describes how you are feeling by circling a number on the scale. "1" means "does not apply at all", and "7" means "applies very much" to how you are feeling. Don't spend much time thinking about each word, just give a quick, gut-level response. It is important that you respond openly and honestly.

	applies very much					does not apply at all	
1. bothered	1	2	3	4	5	6	7
2. angry at myself	1	2	3	4	5	6	7
3. friendly	1	2	3	4	5	6	7
4. guilty	1	2	3	4	5	6	7
5. pleased with myself	1	2	3	4	5	6	7
6. irresponsible	1	2	3	4	5	6	7
7. uneasy	1	2	3	4	5	6	7
8. depressed	1	2	3	4	5	6	7

9. happy	1	2	3	4	5	6	7
10. embarrassed	1	2	3	4	5	6	7
11. negative	1	2	3	4	5	6	7
12. concerned	1	2	3	4	5	6	7
13. frustrated	1	2	3	4	5	6	7
14. annoyed at myself	1	2	3	4	5	6	7
15. energetic	1	2	3	4	5	6	7
16. good about myself	1	2	3	4	5	6	7
17. irritated at others	1	2	3	4	5	6	7
18. disappointed with myself	1	2	3	4	5	6	7
19. tense	1	2	3	4	5	6	7
20. disgusted with myself	1	2	3	4	5	6	7
21. regretful	1	2	3	4	5	6	7
22. optimistic	1	2	3	4	5	6	7
23. disgusted with others	1	2	3	4	5	6	7
24. content	1	2	3	4	5	6	7
25. distressed	1	2	3	4	5	6	7
26. uncomfortable	1	2	3	4	5	6	7
27. sad	1	2	3	4	5	6	7
28. helpless	1	2	3	4	5	6	7
29. ashamed	1	2	3	4	5	6	7
30. satisfied with myself	1	2	3	4	5	6	7
31. angry at others	1	2	3	4	5	6	7
32. self-critical	1	2	3	4	5	6	7
33. good	1	2	3	4	5	6	7
34. anxious	1	2	3	4	5	6	7
35. proud	1	2	3	4	5	6	7

**** Please be sure you completed every item. ****

In the preceding questionnaire, people often report that they are experiencing different types of feelings. The next three questions concern why you were experiencing the various feelings.

1. On the previous page you may have indicated that you were feeling somewhat negative about yourself (e.g., guilty, regretful, self-critical, disappointed with myself, etc.). If you had negative feelings about yourself, please explain WHY you were feeling that way.

2. You may have also indicated that you were feeling somewhat positive (e.g., optimistic, happy, content, good, etc.). If you had positive feelings, please explain WHY you were feeling that way.

3. You may also have been feeling somewhat uncomfortable (e.g., anxious, tense, bothered, uneasy, etc.). If you were feeling uncomfortable, please explain WHY you were feeling that way.

APPENDIX E

JOKE RATING TASK

JOKES PRESENTED TO PARTICIPANTS:

Practice joke 1: What do you call a dog with no legs?

Nothing . . . because if you called him, he couldn't come anyhow.

Practice joke 2: Why don't cannibals eat clowns?

They taste funny.

1. How do you get 50 Haitians into a shoebox?

Tell them it floats.

2. Why did the gay divorce his wife?

Because she wouldn't roll over and take it like a man.

3. Did you hear that 37 people were killed in an auto accident in Los Angeles?

Two carring carrying Mexicans collided.

4. What's a Korean seven-course meal?

Six puppies and a bag of rice.

5. Why is Stevie Wonder always smiling?

He doesn't know he's black.

6. How do you know you're old?

When "getting a little action" means your prune juice is working.

7. Two gays were in an elevator when it stopped, the door opened, and a man standing outside asked, "Going down?"

One of the gays said, "Of course not. We're just talking".

8. What's the difference between American Airlines and Mexican Airlines?

On American, a steward offers you a drink for \$4.00; on Mexican Airlines, a steward offers you his sister for \$4.00.

9. How many feminists does it take to change a lightbulb?

Three. One to screw it in, one to lecture about how the socket is exploited, and one to secretly wish she was the socket.

10. Leroy is a 20 year old 9th grader who must use this word in a sentence: distress.

"Distress . . . Girl, put on distress and let's go, I be hungry".

11. What's the definition of panic?
The feeling you get when your plane lifts off the runway and the pilot announces, "Good afternoon, my name is Mohammed . . .".
12. What's an optimist?
A guy who gets married at age 87 and looks for a house close to a school.
13. What's a Mexican fortune cookie?
A taco with a food stamp.
14. How can you tell a gay bar?
The entrance is in the rear.
15. What's the most confusing holiday for black kids?
Father's Day.
16. How many Japanese industrialists does it take to change a lightbulb?
Six. One to screw it in, and five to plan how to dump the old bulb on the American market.
17. What do you do when a dishwasher stops working?
Slap her and tell her to get back to work.
18. What's the most useless thing in Grandma's house?
Grandpa's thing.

APPENDIX F

BEHAVIORAL INTENTION (PHONE CALL) MEASURE

Student Orientation, Advising, and Registration (SOAR)

*Student Orientation Programs
Gordon Commons Offices
717 W. Johnson
Madison, WI 53706-1790*

"REACH OUT" VOLUNTEERS**(SOAR 1998)**

The SOAR program coordinators are looking for volunteers to "reach out" to new students shortly after they participate in SOAR 1998. The primary goal of the Reach Out program is to provide new students the chance to talk with a peer, one-on-one, about the "student life" and "diversity" material presented at SOAR.

As one part of the Reach Out program, we are seeking volunteers to make brief phone calls to new UW students to tell them about the resources available on campus to help them treat African American and other minority students fairly. We expect these phone calls to take about 5-10 minutes each. Your phone calls will tell new students about specific resources available on campus, like the Interim Multicultural Center (IMCC), which are designed to promote diversity efforts. Your calls will also give new students a chance to ask questions and discuss general student life and diversity issues in a more personal manner than they could during SOAR.

If you choose to volunteer, a SOAR coordinator will contact you later this semester with details about what to say during the phone calls. The SOAR coordinator will also give you all the information you'll need about available resources.

The next SOAR sessions will be conducted in April and May 1998. We are seeking volunteers to make phone calls to recent SOAR participants during the last few weeks of the spring semester (beginning around April 15, 1998). Please only volunteer the number of calls you are genuinely willing to make during that time.

_____ Yes, I would be willing to make _____ phone calls.
(Please write in the number of calls you are willing to make.)

_____ No, I do not have time to volunteer.

If you are willing to make phone calls, please provide the following information so a SOAR coordinator can contact you with more details:

Name (please print): _____

Local address: _____

Local phone number: _____