

Trait Inferences as a Function of Automatically Activated Racial Attitudes and Motivation to Control Prejudiced Reactions

Michael A. Olson and Russell H. Fazio

*Department of Psychology
Ohio State University*

This study investigated how automatically activated racial attitudes and motivation to control prejudiced reactions contribute to the impressions formed of targets whose photos varied by race, gender, and occupation. In earlier sessions, participants completed Dunton and Fazio's (1997) Motivation to Control Prejudiced Reactions scale, and underwent a priming procedure (Fazio, Jackson, Dunton, & Williams, 1995) that provided an unobtrusive estimate of their automatically activated racial attitudes. In the final session, participants provided trait ratings of a number of target photos. Automatically activated racial attitudes were related to the trait inferences participants made of Blacks compared to matched Whites. However, this effect was moderated by motivation to control prejudiced reactions. Among the more motivated, both those with more negative and those with more positive attitudes displayed evidence of correction for their attitudes. This discussion focuses on the ways in which different kinds of motivation to control prejudiced reactions may manifest differently.

First impressions are often formed as a function of the category of which the social target is a member. The general lesson from research in impression formation is that categorical-based impressions are relatively effortless and occur under conditions of low motivation or ability, and individuation is effortful and occurs under conditions of high motivation or ability (Brewer, 1988; Devine & Monteith, 1999; Fazio, 1990; Fiske, Lin, & Neuberg, 1999; Fiske & Neuberg, 1990; Kunda & Thagard, 1996; but see Gilbert & Hixon, 1991; Spencer, Fein, Wolfe, Hodgson, & Dunn, 1998). Most models of impression formation assume that stereotypes and other group level information are activated from memory on perception of a group member (e.g., Banaji & Hardin, 1996; Banaji, Harden, & Rothman, 1993; Brewer, 1988; Devine, 1989; Fiske et al., 1999).¹ More

deliberate and careful processing, on the other hand, can occur as the target's power over the perceiver increases (Fiske et al., 1999), under expectations of future interaction with the target (Johnston, Hewstone, Pendry, & Frankish, 1994), increased accountability for one's judgments (Tetlock, 1992), fear of invalidity (Kruglanski, 1989), and with the presence of unexpected, difficult, or surprising information that begs an explanation (Kintsch, 1988; Kunda, 1990).

However, there are some notable qualifications to automatic stereotype activation. Gilbert and Hixon (1991) showed that under conditions of cognitive load, stereotypes may not be activated at all (see also Spencer et al., 1998). Stereotype activation also depends on the goals of the perceiver (e.g., Macrae, Bodenhausen, Milne, Thorn, & Castelli, 1997). Activation can be reduced or eliminated with practice

Requests for reprints should be sent to Michael A. Olson, Department of Psychology, Ohio State University, 1885 Neil Avenue, Columbus, OH 43210–1222. E-mail: olson@psy.ohio-state.edu

¹In the realm of racial prejudice, other models of avoiding prejudiced responses suggest that motivated individuals might “replace,” or “override” prejudiced responses with more positive ones, instead of attempting to correct for the bias associated with their automatically activated attitudes (e.g., Devine & Monteith, 1999; Wilson, Lindsey, & Schooler, 2000). If this were the case, then we would expect to observe a pattern of responding like that of Panel A of Figure 1, when bias is replaced by the same pre-organized nonprejudiced response, regardless of the extent of the original bias. That is,

if more positive responses are replacing prejudiced ones, the judgment should be the same regardless of the extent of the original prejudiced response. If, on the other hand, a “replacement” process were to be operating in Panel B of Figure 1, then the response replacing the prejudiced one would itself have to vary as a function of the extent of the original bias. To the extent that such variable replacement is permitted by such models, the end result is a pattern of responding that would be indistinguishable from the correction process described by Wegener and Petty (1995). Although we find the correction mechanism more plausible, whether correction or replacement is involved is not crucial to the aims of this research.

(e.g., Dasgupta & Greenwald, 2001; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000). Also, Fazio and colleagues (1995) provided evidence that what is automatically activated on encountering a group member is not necessarily a culturally shared stereotype, but one's personal evaluation of the group. That is, the valence of what is automatically activated on encountering a member of a stereotyped group can be negative, positive, or neutral. This research considers the combination of these individual differences in automatically activated racial attitudes and Motivation to Control Prejudiced Reactions on Whites' reported judgments of Blacks in an impression formation setting (Dunton & Fazio, 1997; Fazio et al., 1995).

The MODE model, the theoretical orientation for this research, argues that when one has a strong attitude toward a given object or class of objects, the mere presentation of that object will activate the attitude from memory spontaneously (Fazio, 1990; 1995; Fazio & Towles-Schwen, 1999). Once activated, attitudes can guide attention toward the object (Roskos-Ewoldsen & Fazio, 1992), color perceptions of the object (Fazio, Roskos-Ewoldsen, & Powell, 1994), and direct judgments about and behavior toward the object (see Fazio & Towles-Schwen, 1999, for a review). In the race domain, if one's attitude toward members of a particular group is capable of automatic activation, seeing a member of that group will activate the attitude and may affect judgments and behavior toward the group member (Fazio, 1990; Fazio & Towles-Schwen, 1999).

The MODE model also makes predictions regarding the conditions under which judgments and behavior toward attitude objects will be affected by more deliberate and effortful thinking. Specifically, when both the motivation to deliberate and the opportunity to do so exist, judgments and behavior toward an attitude object are more likely to be influenced by a careful consideration of the known attributes of the object and by the specific goals of the perceiver. However, our judgments of attitude objects are probably only rarely the products of either automatic or motivated processes. Thus, much of the dirty work of the dual-process models has been disentangling the relative contributions of both automatic and motivated processes on our judgments and impressions of others. Thus, the MODE model includes a large role for "mixed" processes, ones that lead to judgmental outcomes stemming from both automatic and controlled components. These are cases when an automatic response to a stimulus serves as an initial input to a judgment, and is either combined with or attenuated by more careful deliberation of the target or the potential decisions (see Gilbert, 1999, for a more thorough overview of the possible ways automatic and controlled processes interact in judgments).

WHENCE MOTIVATION?

The question of when perceivers are likely to engage in more motivated processing can be more broadly construed as a

question of the perceived costs and benefits of increased effort and attention (Fazio & Towles-Schwen, 1999). The MODE model suggests that social perceivers weigh the cost of reaching a particular conclusion about a target, and the extent to which a particular judgment might produce positive or negative outcomes for the individual. Negative outcomes can occur because of a judgment's deviation from either reality or accepted internal or external norms for behavior.

It is important to note that the costs of reaching a particular conclusion about a target can extend beyond issues of accuracy. Some individuals may weigh other factors, such as the possibility of appearing prejudiced, more heavily than accuracy when reporting impressions of social targets. More specific to these purposes, some Whites may report an especially positive impression of a Black target to avoid appearing prejudiced. For example, Carver, Glass, and Katz (1978) exposed participants to transcripts of interviews supposedly conducted with Black and White undergraduates in a first impression situation. They found that Whites reported especially positive impressions of the Black target on traditional explicit measures, but reported relatively negative impressions of Blacks when connected to a bogus pipeline, which could supposedly measure their true attitudes. Carver et al.'s participants apparently experienced a motivation to avoid prejudiced responding.

Similar findings indicative of preferential evaluations of Blacks on verbal measures have been reported by other researchers (Biernat & Vescio, 1993; Biernat, Vescio, & Theno, 1996; Gaertner & Dovidio, 1977; Jussim, Coleman, & Lerch, 1987). Arguably, this active preferential treatment of Black targets is due to a desire to appear nonprejudiced, or at least to adhere to norms prohibiting prejudice. Yet, whereas the research just described documents that Whites sometime display preferential treatment of Blacks, this research seeks to uncover *who* is most likely to "bend over backwards" in reporting evaluations of Black target persons.

AUTOMATICALLY ACTIVATED RACIAL ATTITUDES AND MOTIVATION TO CONTROL PREJUDICED REACTIONS

Traditional measures of racial attitudes are not capable of disentangling the relative contributions of automatic and motivated processes on race-relevant judgments because they themselves can be influenced by both sorts of processes. An unobtrusive priming measure of racial attitudes developed by Fazio and colleagues provides a means of assessing evaluations that are automatically activated on encountering the attitude object. Because participants are both unaware that their attitudes are being measured, and lack the opportunity to evoke controlled processes while completing the measure, Fazio et al.'s (1995) "bona fide pipeline" is believed to assess the automatic component of racial prejudice. (Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Fazio et al., 1995). On a given trial, the participant indicates the connotation

("good" or "bad") of an adjective presented on a computer screen. The latency to respond to the adjective, and how it is affected by a preceding prime, is used to calculate the attitude estimate. In the case of Whites' attitudes toward Blacks, both Black and White faces serve as primes. If negativity is automatically activated on encountering a Black target, being presented with a Black prime should activate negativity and facilitate identification of negative adjectives as negative, and slow identification of positive adjectives as positive. Attitude estimates derived from this technique have proven predictive of a number of important race-related judgments and behaviors across several studies. Participants characterized by more negative attitudes have behaved more inhospitably toward a Black experimenter (Fazio et al., 1995), judged more harshly an essay supposedly written by a Black undergraduate in a study purportedly involving an essay competition (Jackson, 1997), and viewed a Black candidate for an important volunteer position as relatively less qualified than a comparable White candidate (Olson & Fazio, 1999).

The MODE model argues that the impact of automatically activated attitudes on judgments will be diminished insofar as ample opportunity and motivation exist to counter the effects of these attitudes. Opportunity is required because inferences based on more motivated processes require more time and effort than those based on more automatic processes (see Fazio & Dunton, 1997, for relevant latency data documenting this point). In this study, the opportunity factor is held constant and high, and specific kind of motivation, Motivation to Control Prejudiced Reactions (MCPR), is considered (Dunton & Fazio, 1997). The MCPR is a 17-item, two-factor scale designed to assess Whites' motivation to inhibit their expressions of prejudiced judgments and behavior. The first factor is "concern with acting prejudiced," consisting of items such as "I get angry with myself when I have a thought or feeling that might be considered prejudiced." The second factor, "restraint to avoid dispute," consists of items designed to assess the extent to which one is willing to restrain oneself in the interest of avoiding dispute with or about Blacks. It consists of items such as, "If I were participating in a class discussion and a Black student expressed an opinion with which I disagreed, I would be hesitant to express my viewpoint."

The two factors of the MCPR seem to relate differently to several important race-related variables. Regarding sociopolitical orientations, a recent study showed that the concern factor related strongly to humanitarianism-egalitarianism $r = .50$, to which the restraint factor bore no relationship $r = -.01$ (Fazio & Hilden, 2001). In a study on childhood race-related experiences, Towles-Schwen and Fazio (2001b) found that the concern factor correlated highly with the positivity of Whites' experiences with Blacks in elementary and middle school, and correlated negatively with reported levels of parental prejudice. Thus, a parental emphasis on egalitarianism and relatively positive experiences with Blacks were associated with greater concern with acting prejudiced. The restraint factor, on the other hand, was re-

lated to relatively infrequent (and when they did occur), less intimate, and more negative experiences with Blacks in childhood. Higher restraint individuals also reported gaining more knowledge of Blacks from television and media in lieu of direct personal experience, and claimed that their parents were relatively prejudiced.

Concern and Restraint also are associated differently with the emotions experienced after exhibiting a seemingly prejudiced response. In the context of a study ostensibly about "emotional reactions to television commercials," Fazio and Hilden (2001) exposed participants to a public service ad designed to lead viewers to assume wrongly that a particular Black man was a criminal, thus provoking a seemingly prejudiced response. Understandably, both high concern and high restraint individuals felt agitated at being duped by the commercial. However, high concern individuals also felt guilty for their reactions, whereas high restraint individuals did not. Thus, the two factors of the MCPR relate to race-related judgments and emotions, but they do so differently.

More relevant to these purposes, MCPR scores have also been shown to moderate the effects of automatically activated racial attitudes. For example, attitude estimates derived from the bona fide pipeline were found to predict participants' scores on the Modern Racism Scale (MRS; McConahay, 1986), but only among individuals characterized by relatively low motivation to control prejudiced reactions (Fazio et al., 1995). More motivated individuals showed evidence of having corrected for the influence of their attitudes—those characterized by automatically activated negativity described themselves as especially nonprejudiced on the MRS. Similarly, in research reported by Dunton and Fazio (1997), more negative racial attitudes were associated with more negative ratings of "the typical Black male undergraduate" only among low-motivated participants. Thus, in accord with the MODE model, motivation to control prejudiced reactions can moderate the effects of automatically activated attitudes on race-related judgments.

BIAS CORRECTION IN RACE-RELATED JUDGMENTS

Dunton and Fazio (1997) speculated that the process by which motivation to control prejudiced reactions might operate, in conjunction with automatically activated racial attitudes, in determining judgments may follow that depicted by Wegener and Petty's (1995) Flexible Correction Model. It posits that individuals correct for perceived bias, when motivated and able, according to their idiographic naive theories about how a given source of bias might influence their judgments. Their model serves as a guide to our thinking about the corrective measures that motivated individuals to take control prejudiced responding. Of interest, the moderating relations observed by Dunton and Fazio varied somewhat as a function of the specific motivational factor and the type of judgment being made. In predicting participants' MRS

scores, the concern factor produced both a main effect (such that highly concerned individuals reported relatively less prejudice), and an interaction with automatically activated racial attitudes (such that negative attitudes were associated with prejudiced MRS scores only for people characterized by low concern). Restraint factor scores did not relate to MRS scores. When rating the “typical Black male undergraduate,” a similar main effect of the concern factor was observed, but it was the restraint factor, not the concern factor, that interacted with automatically activated attitude estimates in predicting participants’ responses. Moreover, as we shall explicate momentarily, the pattern of the interaction involving the concern factor with respect to the MRS and that involving the restraint factor with respect to the “typical Black male undergraduate” differed in an interesting and informative way.

Hypothetically, the interaction between motivation and automatically activated attitudes may assume one of four relevant forms. As Figure 1 illustrates, bias correction can be conceptualized in terms of its goal and extent. The top two panels illustrate correction in the interest of reaching a posi-

tive judgment and, hence, display greater correction for more negative attitudes. That is, the more their automatically activated attitudes depart from the goal of being positive, the more appropriately motivated individuals correct for those attitudes. Panel A portrays a relationship between motivation and attitudes that, in terms of Wegener and Petty’s (1995) model, implies an approximately appropriate degree of correction for more negative attitudes. The flat line for highly motivated individuals suggests that they were able to avert the influence of their negative racial biases to an extent that resulted in their reported judgments mimicking those of individuals with more positive attitudes. Panel B portrays a motivation by attitude interaction characterized by overcorrection for more negative attitudes. Again, motivated Whites with more negative biases toward Blacks correct for their bias, but in terms of the Flexible Correction Model, these individuals overestimate the extent to which that bias influences their judgments and, hence, overcompensate. That is, they express judgments that are even more positive than those reported by unmotivated individuals with positive attitudes.

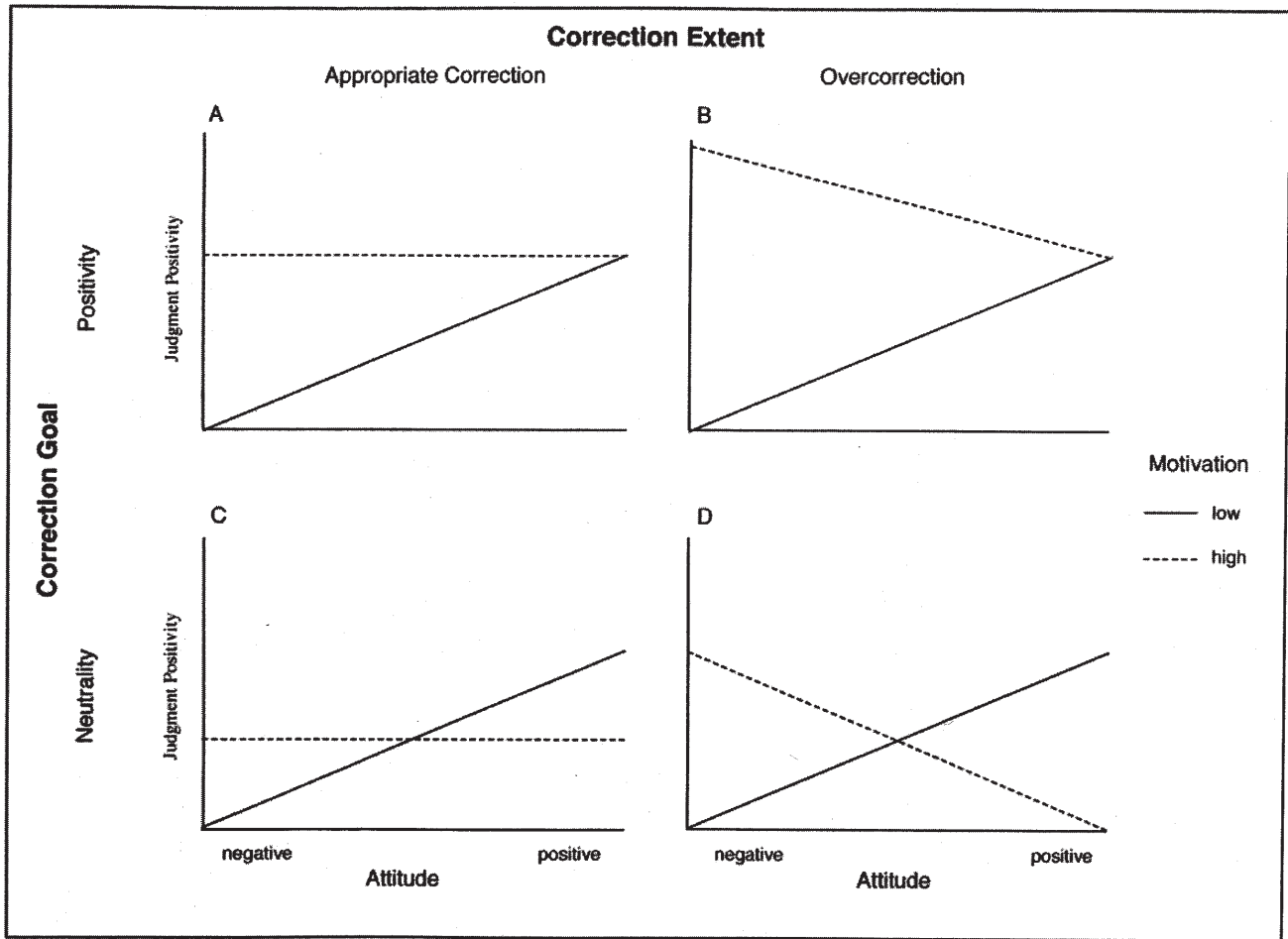


FIGURE 1 Hypothetical forms of attitude × motivation interactions predicting race-related judgments and behaviors.

The bottom two panels (C and D) portray a null and an inverse relationship, respectively, between attitudes and judgments for the more motivated, just as was true for Panels A and B. What differs is the point of intersection with the line for the less motivated—a point that is located near the neutral value of the attitude scale. These scenarios involve correction in the interest of reaching a judgment that is impartial with respect to race. The goal here is not to respond as someone with a positive attitude might, but to respond as someone with a neutral attitude might. In other words, the intent is to be “color-blind” (Frankenberg, 1993). As a result, people with both negative and positive racial attitudes are attempting to compensate for their automatically activated attitudes. Panel C illustrates appropriate correction, whereas Panel D illustrates overcorrection. Comparing Panels B and D, it is apparent that an inverse relationship between attitudes and judgments among the more motivated can stem either from overcorrection in the interest of reaching a positive judgment by those with more negative attitudes, or from overcorrection in the interest of reaching a judgment that is neutral with respect to race by those with both positive and negative attitudes.²

Let us now return to Dunton and Fazio’s (1997) findings. Of interest, the moderating effects of concern on the MRS resembled Panel B, implying overcorrection for negativity. On the other hand, the moderating effects of restraint on ratings of the typical Black male undergraduate resembled Panel D, implying overcorrection for bias among motivated individuals with both negative and positive racial attitudes. Thus, inverse relations between attitudes and judgments among the more motivated participants were observed for both the Modern Racism Scale and the evaluations of the typical Black male undergraduate. It was these inverse relations that prompted Dunton and Fazio to speculate that individuals with stronger motivation to control prejudiced reactions might engage in theory-based correction processes like those described by Wegener and Petty. However, Dunton and Fazio (1997) failed to recognize the potential significance of the two different patterns—concern promoting correction only for negative bias, versus restraint promoting correction for both negative and positive bias.

Why might the different patterns emerge? As noted earlier, the concern with acting prejudiced factor is strongly related to egalitarianism. Such an egalitarian orientation implies attentiveness toward negative biases against historically disadvantaged groups. Given that their goal is to treat such disadvantaged individuals more favorably, those motivated by egalitarian values may be inclined toward positive judgments and, hence, may correct for any negativity that they experience. On the other hand, individuals with more positive attitudes believe they have nothing to correct for, that their automatically activated positivity toward Blacks is concor-

dant with their goals. Thus, it seems appropriate to characterize concern with acting prejudiced as a unidirectional force that promotes correction for negativity in the interest of treating Blacks positively (as in panel B).

In contrast, a motivation to exercise restraint in the interest of avoiding dispute with or about Blacks has a more bidirectional character. High restraint individuals wish to avoid race-related conflict, and conflict may result not only from negative bias (e.g., in the form of accusations of prejudice), but positive bias as well (e.g., in the form of accusations of “bending over backwards” for Blacks, or “reverse-discrimination”). Whites with positive attitudes toward Blacks are not immune to race-related dispute and, in the interest of avoiding such dispute, they may correct for their positivity, much in the same way that similarly motivated individuals may correct for their negativity (as in panel D).

The possibility that some people with positive biases toward Blacks might be motivated to correct for that bias is worth examining more closely. Not only is the possibility somewhat counterintuitive, but it appears inconsistent with some findings in the literature that have documented a greater likelihood of correction for negative information than for positive information when judging a target person. For example, Lambert, Khan, Lickel, and Fricke (1997) presented evidence that perceivers tend to view the negative implications of a stereotype as a more inappropriate basis for judgment than the positive implications and, hence, correct more for negative stereotypes than for positive ones. Similarly, Wyer and Budesheim (1987) found greater adjustment when participants were instructed to disregard previously presented unfavorable information about a target than when they were to disregard favorable information.

Although we do not question the existence of such asymmetries, we also believe that the motivational force of restraint to avoid dispute may represent a special case in that the desire to avoid dispute with or about Blacks may characterize individuals with positive attitudes, as well as those with negative attitudes. At this point, however, the only support for this bidirectional hypothesis is Dunton and Fazio’s single finding regarding evaluations of the typical Black male undergraduate. One might question whether this is a reliable effect, especially in light of the asymmetry suggested by other research. One of the goals of this study, therefore, was to examine whether bidirectional correction occurs across a broad selection of targets in a variety of contexts—essentially, whether it can be referred to as a more general phenomena. Pictorial stimuli analogous to the race-gender-role form of “Black male undergraduate” were employed. More specifically, White participants were shown photos of Black and White men and women that also included occupation information. The photos were selected from a previous study in which participants made similarity ratings of numerous pairs of photos (Fazio & Dunton, 1997).

We expected automatically activated racial attitudes and Motivation to Control Prejudiced Reactions to contribute to

²See footnote 1.

the impressions Whites form of Black targets in accordance with the MODE model's predictions: automatically activated racial attitudes should relate more positively to judgments as motivation decreases. More specifically, however, and based on Dunton and Fazio's (1997) findings, we expected the restraint factor of the Motivation to Control Prejudiced Reactions scale to moderate the effects of attitudes on impressions. And, based on our reasoning regarding the bidirectional nature of restraint to avoid dispute, we also expected that individuals with higher restraint scores would display judgments suggestive of correction for both automatically activated positivity and automatically activated negativity. Among those with negative attitudes, higher restraint individuals were expected to "bend over backwards" to reach favorable judgments of the Black targets. Higher restraint individuals with positive attitudes were expected to compensate for their positivity by expressing relatively less favorable impressions than low restraint individuals.

METHOD

Session 1

In an initial mass survey, several hundred undergraduates in psychology courses completed the Motivation to Control Prejudiced Reactions (MCPR), which was embedded in a number of other scales, for course credit. The full scale, along with its factor structure and additional psychometric properties, can be found in Dunton and Fazio (1997).

Session 2

The second session involved the unobtrusive estimate of participants' automatically activated attitudes toward Blacks using the bona fide pipeline, a priming procedure that consists of 5 separate phases (the fourth of which is the actual priming task, Fazio et al., 1995). One hundred and forty-eight participants who had completed the mass survey were contacted and agreed to participate in the experiment for monetary compensation. They were unaware of any relationship between the first and second sessions. On entering the lab, participants were greeted by a White female experimenter and seated individually in cubicles containing a high-resolution color monitor. They were told the experiment was about word meaning as an automatic skill.

The first phase of the procedure is designed to obtain individual base-line response time data. Participants were instructed to respond to adjectives of either a positive or negative valence appearing on the screen by pressing a key labeled "good" or "bad" on a customized response box. The positive adjective list consisted of 12 items such as "attractive" and "likeable," and the negative adjective list included 12 items such as "disgusting" and "offensive." For each trial, the adjective was preceded by a string of asterisks as a warn-

ing that an adjective would soon appear. Each adjective remained on the screen until a participant responded, or for a maximum of 1.75 sec. Each trial was separated by 2.5 sec. Responses and latencies to respond (from adjective onset to response) were recorded. Participants were instructed to respond as quickly and accurately as possible. There were two blocks of 24 trials each, and a practice block preceded the critical blocks. Order of presentation of the adjectives was randomized for each block. The average latency of response to an adjective across the two blocks served as a participant's baseline for that adjective.

Phases 2 and 3 were designed to prepare participants for the priming phase and obscure our interest in race. Phase 2 ostensibly consisted of a face-learning task, where participants were instructed to simply attend to a number of faces (16 black and white yearbook photos of Black, White, and Asian faces) that appeared on the screen for a subsequent recall task. Each photograph was presented twice—once for each of two blocks. Phase 3 involved a recognition test for the faces presented in Phase 2, where participants indicated whether they recognized the faces presented to them by pressing keys labeled "yes" and "no." Thirty-two faces were presented, 16 of which were new. Faces remained on the screen for a maximum of 5 sec each, or until the participant responded. The intertrial interval was 2.5 sec.

Participants were told that Phase 4 (the actual priming phase) involved combining Phases 1 and 2 under the purported argument that if word meaning identification were truly an automatic skill, the addition of a face learning task should not inhibit performance in identifying the valence of adjectives presented on the screen. Participants were told to attend to the faces again because a recognition task would follow. The primes used for Phase 4 were 48, 256 color, 640 × 480 digitized photos of Blacks, Whites, Asians, and Latinos. All were head and shoulder yearbook-style photos of undergraduate students, and none had appeared in earlier phases. There were 4 blocks of 48 trials. On each trial, a prime was presented for 315 ms, followed by a 135 ms interval, followed by the target adjective (the same 24 adjectives used in phase 1). Thirty-two of the 48 total trials consisted of pairs of photos matched by adjective and sex but varying the race of the photo, either Black or White. Over the course of the four blocks, each photo prime was presented with two positive and two negative adjectives. Black-White matched pairs of a given gender were constructed by ensuring that the same four adjectives followed each photo of the pair. Additional filler trials with White, Asian, and Latino faces obscured the nature of the experiment.

Phase 5 consisted of the face detection task participants had been led to expect. Their task involved identifying whether 48 color photos of faces were "old" or "new" by pressing an appropriate button (half were old and half were new). Faces were presented for 5 sec, with a 2.5 sec intertrial interval. The instructions were identical to phase three. After completion of the final phase, participants were thanked and dismissed from

the experiment. As participants exited, a different lab worker asked them if they would be interested in participating in one of two additional experiments for course credit. Many agreed to participate in what was to be Session 3.

Session 3

Fifty-nine participants returned between 3 and 5 weeks. Forty were women; all were White. They were told that the experiment was about first impressions, and that they would be viewing a number of photographs of people depicted in various occupations. Their task was to rate each person on 5 trait terms on a 6-point scale ranging from 0 (*not at all*) to 6 (*very much so*). The trait terms were: “intelligent,” “industrious,” “likeable,” “honest,” and “reliable.” Participants were informed that they had 30 sec to complete the 5 trait ratings for each photo. Each photo presentation was separated by 6 sec. The presentation order was randomized for each participant. One female participant failed to complete the dependent measure, and therefore was not included in analyses.

Participants made trait inferences of 32 photos in total, 16 of which were fillers of White and Asian individuals included only to obscure the nature of the task. The critical photos consisted of 8 Black–White pairs that were matched by gender of the person depicted, status of the occupation, and independence of the occupation (see Fazio & Dunton, 1997).³ The female pairs consisted of a Black receptionist matched with a White business woman, a Black nurse matched with a White pharmacist, a Black repair woman matched with a White painter, and a Black cashier matched with a White police officer. The male pairs consisted of a Black minister matched with a White professor, a Black businessman matched with a White architect, a Black gardener matched with a White sanitation worker, and a Black potter matched with a White brick layer. After completing the trait inferences of all 32 photos, participants were debriefed, thanked, and dismissed. No participant reported any suspicions about the connections between the sessions.

RESULTS

Preliminary Analyses

Motivation to control prejudiced reactions scale.

Participants had completed the Motivation to Control Prejudiced Reactions scale in the first session. Replicating previous studies using the scale, the two factors, “Concern for

³Multidimensional scaling revealed the similarity judgments to be a function of four underlying dimensions: race, gender, status of the occupation, and the independence of the occupation (i.e., the extent to which it involved working alone or working with others). The multidimensional space generated from this previous study allowed us to select photos that varied by race, but were matched, not only in terms of their gender, but also in terms of the status and independence of the occupation depicted in the photos.

Acting Prejudice,” and “Restraint to Avoid Dispute,” emerged in a principal components analysis (using varimax rotation) of all the mass survey participants. The factor score coefficients from this analysis were used to compute the two factor scores for each of the participants in this study. The two factors did not correlate with one another, $r(57) = .13, p > .3$. Nor did either factor correlate with the attitude estimate ($r_s < .07, p_s > .60$).

Attitude estimates. Estimates of automatically activated racial attitudes were derived from participants’ performance in the session two priming procedure, just as described in detail in Fazio et al. (1995). Baseline latencies were computed for each participant for each adjective by averaging the latencies of the two presentations of each adjective in phase one. This score was then subtracted from the latency for each adjective’s presentation during the priming phase (phase 4), yielding a facilitation score for each face-adjective combination. Mean facilitation scores on the 2 positive and 2 negative adjectives were then computed for each face (incorrect responses to the adjective connotation were dropped). The 32 critical facilitation scores, 16 for Black faces and 16 for White faces, remain for each participant.

In calculating the attitude index, a Race of Photo \times Valence of Adjective interaction was computed in an ANOVA predicting the facilitation scores for each participant. The effect size of this interaction serves as the attitude index. Negative scores indicate relatively greater facilitation in responding to negative adjectives preceded by Black faces and relatively less facilitation for positive adjectives preceded by Black faces, whereas the reverse is true in the case of positive scores. The attitude estimates ranged from -0.46 to 0.36 , with a mean of -0.03 ($SD = 0.19$), which did not differ from 0, $t(57) = 0.98, p = .32, d = .25$.⁴

Trait Inferences

We first examined relationships among the 5 trait inferences for each Black–White pair (intelligent, industrious, likeable, honest, and reliable). The trait ratings were highly correlated and reliability analyses yielded coefficient alphas ranging from .71 to .87 across the target photos. Thus, we decided to combine the 5 traits into a single mean for each photo for each participant, with positive values indicating more positive trait inferences.

⁴In the many studies that the laboratory has conducted using the bona fide pipeline priming procedure, the average scores sometimes have been significantly more negative than zero (e.g., Fazio et al., 1995; Olson & Fazio, 1999; Towles-Schwen & Fazio, 2001b), and sometimes have not (e.g., Fazio & Dunton, 1997; Fazio & Hilden, 2001; Jackson, 1997; Towles-Schwen & Fazio, 2001a). We presume that these outcomes simply reflect sampling variability. More important, relations between the attitude estimate and race-related judgments and behavior have been observed regardless of the samples’ average level of negativity toward Blacks.

Within-subject correlations between trait ratings and the occupational variables of the photos indicated that participants were attentive to the occupational information in the photos. The mean correlation between occupational status and trait ratings was substantial, $M = 0.67$ ($SD = .19$), $t(57) = 26.47$, $p < .001$, $d = 7.01$, indicating that participants made more positive trait inferences of photos depicting higher status occupations. There is little reason to expect that the independence of an occupation would be related to positive trait ratings, and accordingly, the mean correlation between the two variables was essentially zero, $M = 0.04$ ($SD = .18$), $t(57) = 1.44$, $p = .16$, $d = .38$. So it appears that participants took the task seriously and attended to pertinent information within the photos in arriving at their trait judgments. Participants were also attentive to gender of the individuals depicted in the photos, and made more positive trait inferences of female targets ($M = 4.44$, $SD = 0.53$) compared to male targets ($M = 4.09$, $SD = 0.53$), $t(57) = 6.80$, $p < .001$, $d = 1.80$.

The most intriguing result of our initial analyses was the difference between participants' ratings of Black and White targets. Regardless of status and gender, participants made more positive trait ratings of the Black targets ($M = 4.36$, $SD = 0.52$) than the White targets ($M = 4.17$, $SD = 0.49$). This difference was highly significant, $t(57) = 5.98$, $p < .001$, $d = 1.58$,⁵ and is reminiscent of other reports of Whites showing preferential treatment of Blacks (e.g., Biernat & Vescio, 1993; Biernat et al., 1996; Carver et al., 1978; Jussim et al., 1987).

Predicting Race Differences in Trait Inferences

Our major aim was to examine how such ratings would relate to automatically activated racial attitudes and motivation to control prejudiced reactions. Accordingly, we conducted a mixed-design multiple-regression analysis predicting the average ratings of the Black versus White stimulus persons from the race of the photo (a within-subjects factor), and from three between-subjects factors—the estimates of racial attitudes based on the priming procedure, and the scores on the two motivation factors (Cohen & Cohen, 1983). The between-subjects portion of this analysis revealed a main effect of the concern factor of the MCPR, $t(54) = 3.06$, $p < .01$, $b = .39$, $SE = .069$, $d = .83$. In general, participants with higher concern scores made more positive trait inferences of the targets. This effect was unqualified by race of the target. Thus, higher concern was associated with indiscriminate positivity toward the targets regardless of race. None of the other between-subjects variables produced main effects, nor were any interactions among the between-subjects variables apparent all $t_s < .8$, $p_s > .4$, $d_s < .2$.

⁵A 2 (race of photo) \times 2 (sex of photo) repeated measures ANOVA was conducted on the mean trait inferences. Results matched those of the t tests—both main effects were highly significant ($p_s < .001$), and the Race \times Sex interaction was not, $F(1, 56) < .86$, $p > .6$, $d = .24$.

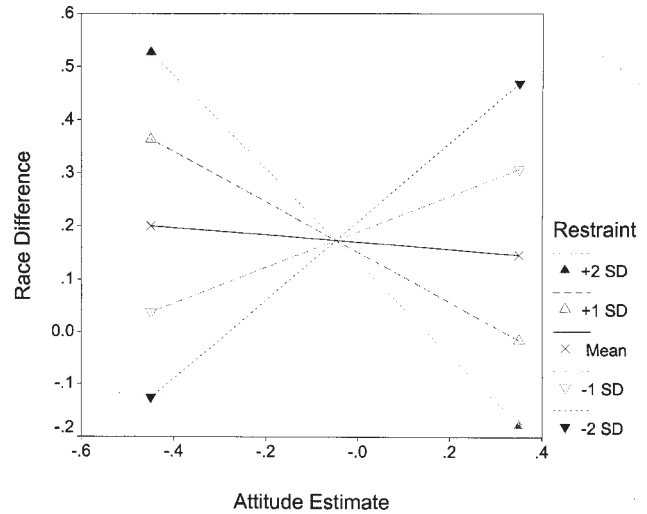


FIGURE 2 Trait inferences as a function of automatically-activated racial attitudes at various values of the restraint factor of the MCPR. Higher numbers indicate more positive trait inferences of Black relative to White targets.

Examination of the within-subjects variance revealed a main effect of race of the photo, thus duplicating the effect of race observed in the preliminary analyses. However, the effect was qualified by the predicted three-way interaction between race of photo, attitude, and restraint, $t(51) = 2.28$, $p = .027$, $b = .30$, $SE = .18$, $d = .64$.⁶ No other effects were significant all $t_s < .7$, $p_s > .5$, $d_s < .2$.

For our purposes, the ratings of the White targets provide a baseline comparison that controls for participants' use of the trait scales. Hence, the nature of the Race \times Attitude \times Restraint interaction can be communicated effectively by considering the difference between the trait ratings of the Black targets and those of the White targets as a function of attitude and restraint scores. Figure 2 displays the family of regression lines indicated by the significant interaction. The relation between racial attitudes and the difference scores is depicted for the mean restraint factor score, for values plus or minus one standard deviation beyond the mean, and for values plus or minus two standard deviations (Aiken & West, 1991). At the mean restraint score, the regression weight is essentially zero ($b = -.07$, $SE = .17$) and, hence, the line is virtually flat. As restraint decreases, the regression weight grows increasingly more positive. At one standard deviation less than the mean, the

⁶The only other effect to emerge in the regression analysis was an uninterpretable interaction involving race of photo, concern, and restraint, $t(51) = 2.51$, $p = .015$, $d = .70$. Relative preference for the Black targets was more evident for participants with corresponding scores on the concern and restraint factors than for participants characterized by high concern and low restraint or low concern and high restraint. Most important, from our perspective, the critical Race \times Attitude \times Restraint interaction on which the text focuses was not itself qualified by the concern factor. That is, no four-way interaction was apparent, $t(49) = .61$, $p = .54$, $d = .17$.

regression coefficient of .34 ($SE = .23$) is marginally significant, $t(51) = 1.47$, $p = .074$, $d = .41$. At two standard deviations less than the mean, the regression coefficient is .74 ($SE = .26$), which differs significantly from zero, $t(51) = 1.98$, $p = .026$, $d = .55$. Thus, for individuals with relatively low restraint scores, the more positive their automatically activated attitudes, the more favorable the impressions they formed of the Black targets relative to the White targets.

As restraint increases over the mean value, the regression weight grows increasingly negative. At one standard deviation greater than the mean, the coefficient is $-.47$ ($SE = .27$), $t(51) = 1.84$, $p = .036$, $d = .52$. At two standard deviations greater than the mean, it assumes a value of $-.88$ ($SE = .31$), $t(51) = 2.15$, $p = .018$, $d = .60$. Thus, for individuals with relatively high scores on the restraint factor, the more positive their automatically activated racial attitudes, the less favorable their impressions of the Black targets. That is, the relation is the inverse of what was observed for the participants characterized by lower restraint.

DISCUSSION

On the average, our White participants reported more positive impressions of Black targets relative to White targets. This finding is similar to occasional reports of Whites “bending over backwards” to favor Blacks (Biernat & Vescio, 1993; Biernat et al., 1996; Carver et al., 1978; Gaertner & Dovidio, 1977; Jussim et al., 1987). This study offers insight concerning the meaning of such preferential treatment, as well as the role that automatically activated racial attitudes and motivation to control prejudiced reactions play in determining its extent. These findings proved supportive of the MODE model’s general predictions, and of our more specific theoretical reasoning regarding the bidirectional nature of the correction forces prompted by restraint to avoid dispute with or about Blacks. The interaction that was observed closely resembled that of Panel D in Figure 1. For participants characterized by relatively low scores on the restraint factor of the MCPR, automatically activated racial attitudes were predictive of the impressions formed of Black targets—more positive attitudes were associated with more positive impressions of Blacks. However, motivated individuals (in this case, people scoring high on the restraint factor) displayed the reverse relation—a pattern that, as we argued earlier, suggests they overcorrected for their automatically activated racial attitudes in the impressions they reported of the Black targets. That is, high restraint individuals for whom negativity was activated reported impressions of Blacks that were just as favorable as those exhibited by low restraint individuals with positive attitudes. Similarly, high restraint individuals for whom positivity was activated, reported evaluations that did not represent much favoritism toward Blacks, and did so

roughly to the same extent as unmotivated (low restraint) individuals with negative attitudes.⁷

In considering the findings for the participants characterized by relatively high restraint scores, it is useful to focus on the potential costs associated with making a prejudiced, a strictly egalitarian, or an “antiprejudiced” (actively favoring Blacks) response. The desire to avoid dispute appears capable of guiding differential evaluations of Blacks versus Whites either in the direction of greater positivity toward Blacks or in the direction of negativity toward Blacks. What happens seems to depend on the nature of the attitude that is automatically activated. It is, after all, the influence of this attitude that these individuals are motivated to control. For high restraint individuals with negative attitudes toward Blacks, the worry seems to be that their negativity may provoke dispute, and hence reporting an overly-positive impression of Black targets serves to reduce the likelihood of dispute. That is, they bend over backwards. Effects similar to this have been observed in past studies that revealed a correction process among motivated individuals characterized by the activation of negativity in response to Blacks (e.g., Dunton & Fazio, 1997; Fazio et al., 1995). Perhaps more striking is the indication that high restraint individuals with positive automatically activated racial attitudes also tended to correct for their attitudes. The greatest worry for these individuals might be an *accusation* of “bending over backwards” (i.e., “reverse discrimination”) because of their positivity, and in a similar fashion, reporting a less positive impression of Black targets is the best route to avoid such an accusation. Thus, the nature of the particular motivational force underlying restraint to avoid dispute is capable of evoking corrective processes for both negative and positive racial biases.

Although the limited base of findings indicative of motivated control for prejudiced reactions forces us to be speculative, we note an interesting difference between those studies in which concern with acting prejudiced has moderated the influence of automatically activated racial attitudes versus those in which restraint to avoid dispute has acted as the moderating force. The former has been observed with respect to responses to the Modern Racism Scale, whose scale items explicitly refer to “Blacks” as a category (Dunton & Fazio, 1997). Moderating effects of concern (but not restraint) also were observed by Towles-Schwen and Fazio (2001a) in a

⁷The single finding that seems somewhat inconsistent with the framework we have outlined is Dunton and Fazio’s (1997) observation of a main effect of the concern factor, where individuals with higher concern for acting prejudiced reported more positive evaluations of Blacks. In this study, however, higher concern was associated with reports of indiscriminately positive impressions of all targets, regardless of race. This finding leads us to question the appropriateness of considering the main effect observed in the Dunton and Fazio study, which did not consider White targets, as race-related. The commitment to egalitarian and humanitarian standards that characterize high concern appears to foster a reluctance to ascribe negative traits to individuals for whom little individuating information is available.

study in which participants reported their anticipated comfort interacting with members of various ethnic and social categories, including “a person who is Black,” in a variety of situations. Automatically activated racial attitudes related to willingness to initiate intimate and relatively unscripted interactions with a Black person, only among participants with relatively low scores on the concern factor. Thus, both of the studies in which concern proved to be the moderator forced participants to consider Blacks solely at the categorical level.

In contrast, consideration of both the “typical Black male undergraduate,” as in Dunton and Fazio (1997), and the traits likely to be exhibited by targets for whom race, gender, and occupation cues are specified, as in this study, showed automatically activated racial attitudes to be moderated by restraint to avoid dispute, and not concern. Such judgments appear to be construed as race-related, but the emphasis is not on the perception of a category. Instead, the issue is how a particular individual, albeit a member of the category, is perceived.

Concern for acting prejudiced, in being closely aligned with egalitarianism, appears to evoke corrective processes when “Blacks” as a group are considered more abstractly and from a distance. It could be that high concern individuals’ orientation toward egalitarianism limits their bias correction to cases where treatment of Blacks is operative at the group level, especially where discrimination against Blacks has received more attention historically (e.g., equal rights, affirmative action). Thus, it appears that any amount of personal information above and beyond the category pulls the social target out of the politics of egalitarian treatment of historically disadvantaged social categories, where moderating effects of concern are observed, and into the realm of interpersonal interaction, where moderating effects of restraint are observed. Putting a face on “Black,” as personalizing information does, may be the catalyst underlying the evocation of restraint, as opposed to concern, because it is within the realm of the interpersonal that dispute is most likely to arise. Speculation aside, questions regarding the contexts that activate one form of motivation over another (not to mention when any motivation to control prejudiced reactions is evoked at all), are important questions for future empirical work.

CONCLUSIONS

In summary, the MODE model’s predictions as to the circumstances under which automatically activated attitudes will guide social judgments were confirmed. Whites made trait inferences about Blacks according to their racial attitude if they were low in restraint to avoid dispute, but corrected for their attitudes if they were high in such motivation. In addition, the findings provide further evidence of the predictive validity of the attitude estimate yielded by the bona fide pipeline priming measure, at least among individuals with low

motivation to control prejudiced reactions (Dunton & Fazio, 1997; Fazio et al., 1995).

Finally, these findings also illustrate the importance of considering motivated processes when examining the predictive validity of an implicit measure of racial attitudes. If we had not assessed motivation to control prejudiced reactions, the data would have led us to the erroneous conclusion that the implicit measure bore no relation to the impressions Whites formed of Black targets. The relation clearly was moderated by motivation, and the moderating influence yielded relations of opposing directions. For individuals low in restraint to avoid dispute, more positive automatically activated racial attitudes were associated with more favorable first impressions of Blacks. Yet, this relation was actually reversed by the motivated correction process exhibited by individuals high in restraint to avoid dispute. Thus, future research concerned with the validity of implicit attitude measures may benefit by considering more automatic and more motivated processes jointly.

ACKNOWLEDGMENTS

As of August 2004, Michael Olson will be affiliated with the University of Tennessee.

This research was supported by Grant MH38832 from the National Institute of Mental Health. Portions of this research were presented at the 1999 Annual Meeting of the Midwestern Psychological Association. We thank Laura Hilden for her assistance in data collection, Suzanne Miller for many helpful discussions about this research, and Duane Wegener for helpful feedback on an earlier version of this article.

REFERENCES

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Banaji, M. R., & Hardin, C. D. (1996). Automatic gender stereotyping. *Psychological Science*, 7, 136–141.
- Banaji, M. R., Hardin, C. D., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology*, 65, 136–141.
- Biernat, M., & Vescio, T. K. (1993). Categorization and stereotyping: Effects of group context on memory and social judgment. *Journal of Experimental Social Psychology*, 29, 166–202.
- Biernat, M., Vescio, T. K., & Theno, S. A. (1996). Violating American values: A “value congruence” approach to understanding outgroup attitudes. *Journal of Experimental Social Psychology*, 32, 387–410.
- Brewer, M. B. (1988). A dual process model of impression formation. In T.K. Krull & R.S. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Carver, C. S., Glass, D. C., & Katz, I. (1978). Favorable evaluations of Blacks and the handicapped: Positive prejudice, unconscious denial, or social desirability? *Journal of Applied Social Psychology*, 8, 97–106.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression: Correlational analyses for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

- Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81*, 800–814.
- Devine, P. G. (1989). Stereotypes and prejudice: Automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5–17.
- Devine, P. G., & Monteith, M. J. (1999). Automaticity and control in stereotyping. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 339–360). New York: Guilford.
- Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23*, 316–326.
- Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75–107). New York: Academic.
- Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Fazio R. H., & Hilden, L. (2001). Emotional reactions to a seemingly prejudiced response. *Personality and Social Psychology Bulletin, 27*, 538–549.
- Fazio, R. H., & Dunton, B. C. (1997). Categorization by race: The impact of automatic and controlled components of racial prejudice. *Journal of Experimental Social Psychology, 33*, 451–470.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013–1027.
- Fazio, R. H., Roskos-Ewoldsen, D. R., & Powell, M. C. (1994). Attitudes, perception, and attention. In P. M. Niedenthal & S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 197–216). San Diego, CA: Academic.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229–238.
- Fazio R. H., & Towles-Schwen, T. (1999). The MODE model of attitude-behavior processes. In S. Chaiken & Y. Trope (Eds.), *Dual Process Theories in Social Psychology* (pp. 97–116). New York: Guilford.
- Fiske, S. T. & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 32, pp. 1–74). New York: Academic.
- Fiske, S. T., Lin, M., & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken & Y. Trope (Eds.), *Dual Process Theories in Social Psychology* (pp. 231–254). New York: Guilford.
- Frankenberg, R. (1993). *White women, race matters: The social construction of whiteness*. Minneapolis: University of Minnesota Press.
- Gaertner, S. L., & Dovidio, J. F. (1977). The subtlety of white racism, arousal, and helping behavior. *Journal of Personality and Social Psychology, 35*, 691–707.
- Gilbert, D. T. (1999). What the mind's not. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 3–11). New York: Guilford.
- Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application stereotypic beliefs. *Journal of Personality and Social Psychology, 60*, 509–517.
- Jackson, J. (1997) Automatically activated racial attitudes. Unpublished doctoral dissertation, Indiana University, Bloomington, IN.
- Johnston, L. C., Hewstone, M., Pendry, L., & Frankish, C. (1994). Cognitive models of stereotype change: Motivational and cognitive influences. *European Journal of Social Psychology, 24*, 237–265.
- Jussim, L., Coleman, L. M., & Lerch, L. (1987). The nature of stereotypes: A comparison and integration of three theories. *Journal of Personality and Social Psychology, 52*, 536–546.
- Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training on the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78*, 871–888.
- Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: Cognitive and motivational issues*. New York: Plenum.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163–182.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480–498.
- Kunda, Z. & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A Parallel-Constraint-Satisfaction Theory. *Psychological Review, 103*, 284–308.
- Lambert, A. J., Khan, S. R., Lickel, B. A., & Fricke, K. (1997). Mood and the correction of positive versus negative stereotypes. *Journal of Personality and Social Psychology, 72*, 1002–1016.
- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., Thorn, T. M. J., & Castelli, L. (1997). On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Social Psychology, 33*, 471–489.
- McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–125). Orlando, FL: Academic.
- Olson, M. A., & Fazio, R. H. (1999). *Nonverbal leakage during public evaluations of Black candidates: The roles of automatically activated racial attitudes and motivation to control prejudiced reactions*. Paper presented at the May Meeting of the Midwestern Psychological Association, Chicago.
- Roskos-Ewoldsen, D. R., & Fazio, R. H. (1992). On the orienting value of attitudes: Attitude accessibility as a determinant of an object's attraction of visual attention. *Journal of Personality and Social Psychology, 63*, 198–211.
- Spencer, S. J., Fein, S., Wolfe, C. T., Hodgson, H. L., & Dunn, M. A. (1998). Automatic activation of stereotypes: The role of self-image threat. *Personality and Social Psychology Bulletin, 24*, 1139–1152.
- Tetlock, P. E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 331–376). San Diego, CA: Academic.
- Towles-Schwen, T., & Fazio, R. H. (2001). On the origins of racial attitudes: Correlates of early childhood experiences. *Personality and Social Psychology Bulletin, 27*, 162–175.
- Towles-Schwen, T., & Fazio, R. H. (2003). Choosing social situations: The relation between automatically activated racial attitudes and anticipated comfort interacting with African Americans. *Personality and Social Psychology Bulletin, 29*, 170–182.
- Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology, 68*, 36–51.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107*, 101–126.
- Wyer, R. S., & Budesheim, T. L. (1987). Person memory and judgments: The impact of information that one is told to disregard. *Journal of Personality and Social Psychology, 53*, 14–29.