Handbook of Prejudice, Stereotyping, and Discrimination

Second Edition
Edited by Todd D. Nelson
THE SELF-REGULATION OF PREJUDICE

Margo J. Monteith, Laura R. Parker, and Mason D. Burns
PURDUE UNIVERSITY

All in all, we are forced to conclude that prejudice in a life is more likely than not to arouse some compunction, at least some of the time. It is almost impossible to integrate it consistently with affiliative needs and human values.

Allport, 1954, p. 329

The fact that prejudice is fundamentally incompatible with humanitarian precepts, egalitarian values, and internalized standards, as noted by Allport (1954), has been the cornerstone of many theories concerning prejudice expression and reduction. For instance, according to symbolic and modern racism theories, Whites express negative affect toward Blacks in ways that justify and rationalize their prejudices (e.g., Sears & Henry, 2003; see also Crandall & Eshleman, 2003). Aversive racism theory maintains that Whites often avoid Blacks so that their underlying negativity may remain unacknowledged, or they express their negativity only when it can be justified with nonracial explanations (e.g., Dovidio & Gaertner, 1998). Thus, people appear to be able to live comfortably with their prejudices through rationalization, justification, denial, and just plain avoiding outgroups. All of these tactics contribute to the maintenance of prejudice rather than to its change.

Another possibility is that people recognize and confront their prejudiced tendencies so that they can work toward prejudice reduction. This pathway to change has been the focus of research on the self-regulation of prejudice. Although Allport (1954) and Rokeach (1973) discussed prejudice reduction in ways that touched on self-regulatory processes, this topic began to receive serious attention in the 1990s—shortly after the publication of Devine (1989). We believe there are three interrelated reasons for this. First, as Devine (1989) argued, it had become increasingly normative for people to sincerely embrace low-prejudice attitudes and to regard prejudice as inconsistent with
their self-concepts. Second, implicit bias measures were invented, along with other techniques for assessing subtle prejudiced responses. Using these measures enabled researchers to demonstrate that even self-professed low-prejudice individuals were prone to intergroup biases through automatic processes that involved little conscious awareness or intention. Devine (1989) demonstrated these subtle biases with a rather cumbersome priming method using a tachistoscope, but more readily available and user-friendly methods such as computer priming (e.g., Fazio, Jackson, Dunton, & Williams, 1995) and the Implicit Association Test or IAT (Greenwald, McGhee, & Schwartz, 1998) were developed soon thereafter (see Monteith, Woodcock, & Gulker, 2013). Third, Devine (1989) set the stage for studying processes of self-regulating prejudice by arguing that low-prejudice people for whom stereotypes were automatically activated were embroiled in a process of change; they needed to "break the prejudice habit" by monitoring themselves and responding on the basis of their low-prejudiced beliefs rather than their automatically activated stereotypes. In sum, the combination of increasingly tolerant outgroup attitudes, technological advances in the measurement of persisting subtle biases, and Devine’s (1989) theoretical perspective all provided inroads for detailed analyses of the self-regulation of prejudice.

In this chapter, we first discuss motivations that can encourage people to self-regulate prejudice. We then review strategies that involve the self-regulation of stereotyping and prejudice and their consequences for the reduction of bias. These strategies include motivational and learning processes that facilitate vigilance against and inhibition of biases so that egalitarian goals can be achieved, attempts to suppress and banish stereotypic thoughts from the mind, and establishing if–then plans that can be implemented in contexts where biased responses are possible. Next we address the extent to which self-regulation consumes energy (e.g., cognitive resources), and ways in which this adverse side effect might be avoided. In addition, we discuss the consequences of people’s use of self-regulation during interracial interactions, including possible costs to regulators but also advantages for the quality of interactions. Finally, we consider how confronting others about their biases in interpersonal context might encourage the self-regulation of prejudice.

**SELF-REGULATION**

Self-regulation has been an important construct in psychological theory and research since the beginnings of psychology (James, 1890/1950). Generally speaking, the act of self-regulation involves setting goals and working toward the achievement of those goals. Both daily life and the accomplishment of long-term objectives require the ability to engage in self-regulation. For example, a goal on a certain day may be to avoid eating chocolate, and the long-term goal may be to lose 10 pounds. With respect to prejudice, a daily goal may be to avoid stereotypic thoughts, and the long-term goal may be to live consistently with one’s egalitarian self-image. Central to the process of self-regulation is the exertion of self-control, or attempts by the self to control the self so as to achieve desired outcomes (Mischel, 1996).

Self-regulation has been studied at many levels, from more macroscopic approaches to the neural level. A traditional model for understanding how individuals go about self-regulation is the cybernetic model (Carver & Scheier, 1990). This model includes four phases in a Test Operate Test Exit (TOTE) sequence. The Test phase involves a determination of whether there is a discrepancy between one’s standards and desired
states. In the Operate phase, behaviors are initiated and enacted for reaching one's goal. The Test phase is then performed again to assess progress, and if the desired goal has been met, one may Exit and discontinue self-regulatory efforts. However, if a discrepancy still exists, further self-regulation occurs to adjust and monitor behavior for goal attainment. Although some strategies for self-regulating prejudice that are discussed in this chapter build on the idea of monitoring and adjusting one's responses, other strategies involve different processes. However, what all strategies have in common is the fact that people must be sufficiently motivated to engage in self-regulation for relevant processes to be instigated.

MOTIVATIONS TO SELF-REGULATE PREJUDICED RESPONSES

A variety of factors affect people's motivation and thus likelihood of attempting to regulate their prejudiced tendencies. One critical factor is people's personal attitudes and standards. People who hold low-prejudice attitudes or standards for responding to members of stereotyped groups are personally motivated to try to respond in egalitarian ways (e.g., Devine, 1989). Nonetheless, much research indicates that people with low-prejudice attitudes are prone to automatic activation of stereotypes and evaluative biases that result in prejudiced responses. That is, without conscious bidding or the intent to have biased thoughts, feelings, or actions, stereotypes can be activated and applied when responding to outgroups (see Bargh, 1999; Dasgupta, 2004) and sometimes even in relation to one's ingroup (i.e., in the case of low-status groups, Ashburn-Nardo, Knowles, & Monteith, 2003; Jost, Pelham, & Carvallo, 2002). Thus, low-prejudice individuals, who are identified with standard attitudinal questionnaires (e.g., Whites who score low on the Attitudes Toward Blacks scale, Brigham, 1993), often need to self-regulate their responses in connection with outgroups.

People who hold more prejudiced attitudes also can be motivated to self-regulate prejudice, although their motivation often stems from external sources and norms. For example, the actual or expected presence of a peer or authority figure who is believed to reprove prejudice leads to reductions of prejudiced self-reports and behaviors (Blanchard, Lilly & Vaughn, 1991; Monteith, Deneen, & Tooman, 1996; Plant, Devine, & Brazy, 2003).

Rather than identify people who are motivated to respond without prejudice based on attitudinal questionnaires, motivations can be assessed directly. Plant and Devine (1998) distinguished between internal and external motivations, which vary orthogonally and are assessed with the Internal and External Motivations to Respond Without Prejudice scales (IMS and EMS). The IMS reflects the degree to which one's motivations are based on personally important values and standards linked to one's egalitarian self-concept (e.g., "Because of my personal values, I believe that using stereotypes about Black people is wrong"). The EMS reflects motivations to abide by external pressures and norms, or the standards imposed by other people suggesting that prejudiced responses are inappropriate (e.g., "I attempt to appear nonprejudiced toward Black people in order to avoid disapproval from others"). Along somewhat parallel lines, Dunton and Fazio (1997) devised the Motivation to Control Prejudiced Reactions Scale, which captures internal and external motivations along with a tendency to restrain one's self from expressing bias that could lead to disputes with others.

Considering whether high IMS individuals are simultaneously more or less externally motivated to control prejudice is important. High IMS–low EMS people appear to be
consistently driven to regulate their prejudice and are even able to do so when biases are assessed implicitly (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002; Gonsalkorale, Sherman, Allen, Klauer, & Amodio, 2011). In contrast, high IMS–high EMS people have intentions to regulate prejudice that are less well-entrained (Devine et al., 2002). Finally, low IMS–high EMS people appear to regulate their prejudice only when pressures from others to appear non-prejudiced are salient (Plant & Devine, 1998).

Legault and colleagues (e.g., Legault, Green-Demers, Grant, & Chung, 2007) have taken a similar approach to understanding the motivational underpinnings of self-regulation, basing their work directly on self-determination theory (Deci & Ryan, 1985). These researchers conceptualize motivation to regulate prejudice on a continuum from non-self-determined to fully self-determined. The self-determined motivation to regulate prejudice expands beyond internal and external sources of motivation by differentiating between various sources of internal and external pressure. Non-self-determined regulation reflects a lack of intention to self-regulate, the desire to avoid social sanction, or an anxiety-driven desire to comply with social standards. Self-determined motivations are driven by valued beliefs, an egalitarian self-concept, or by the desire to pursue egalitarian goals as an end in themselves.

People who lack self-determined motivation inconsistently engage in self-regulation, depending on the external pressures present in a situation (Legault et al., 2007). However, as people become more and more self-determined (i.e., moving toward the pursuit of egalitarian goals as an end in themselves) they more proactively and consistently engage in self-regulation. For instance, participants with high levels of self-determined motivation were significantly less likely to apply activated Black stereotypes when rating an ambiguously hostile target, even though they showed the same extent of stereotype activation on a word fragment task as non-self-determined participants (Legault, Green-Demers, & Edie, 2009). In addition, participants with high self-determined motivation demonstrated significantly lower implicit bias on a race IAT (Legault et al., 2007) and showed resilience to manipulations that increase prejudice and discrimination among non-self-determined participants (Legault & Green-Demers, 2011).

In sum, given sufficient motivation people will self-regulate their prejudice and appear able to generate non-biased responses, even for implicit measures among some people. Precisely how does this self-regulation occur? What processes are involved?

**THE SELF-REGULATION OF PREJUDICE MODEL**

One method for self-regulating prejudice involves motivational, learning, conflict monitoring, inhibitory, and replacement processes. This method is captured by the Self-Regulation of Prejudice (SRP) model (see Figure 17.1) developed by Monteith and colleagues (Monteith, 1993; Monteith, Ashburn-Nardo, Voils, & Czopp, 2002; for reviews, see Monteith & Mark, 2005; Monteith, Lybarger & Woodcock, 2009; Monteith & Voils, 2001). The model starts with the now well-documented fact that stereotypes and implicit evaluative biases can be automatically activated and used as a basis for responding (Devine, 1989). For people who are motivated not to respond in biased ways, such responses result in the occurrence of a discrepant response. If people become aware that they have engaged in a prejudiced response that is discrepant from their standards, the model specifies a variety of consequences that will be critical to subsequent self-regulatory efforts.
One consequence is the experience of negative self-directed affect (e.g., guilty and disappointed with the self). The idea that discrepancies from important self-standards give rise to such affective reactions is well supported by Higgins’s (1987) self-discrepancy theory and related research. Other consequences can be understood in terms of Gray’s (1982; Gray & McNaughton, 1996) theoretical account of motivation and learning. Gray described the behavioral inhibition system (BIS) as the seat of self-regulation. This system initially functions as a comparator, checking for mismatches between expected and actual events. BIS activity has been linked with neural mechanisms associated with conflict monitoring, one of which is the anterior cingulate cortex (ACC; Amodio, Master, Yee, & Taylor, 2008; Botvinick, Braver, Barch, Carter, & Cohen, 2001). When mismatches are detected, as in the detection of prejudiced responses that conflict with one’s motivation to respond without prejudice, arousal increases and ongoing behavior is briefly interrupted. In other words, behavioral inhibition occurs. The BIS also instigates retrospective reflection, such that the discrepant response is allotted enhanced attention so that indicators of it can be identified, such as features of the situation, the environment, and the self. Such attention results in the identification of stimuli that predict the occurrence of the discrepant response. With these consequences comes the natural development of cues for control, or the building of associations between stimuli that predict the occurrence of a discrepant response, the discrepant response itself, and the negative self-directed affect resulting from awareness of one’s discrepancy.

For example, a White woman (call her Elaine) might find herself clutching her purse as she passes a Black man while walking down the street. She might wonder at her reaction, perhaps recognizing that she had passed several White men on the street without clutching her purse, and realize that her behavior is inconsistent with her personal standards for responding without bias in relation to Black people. Her awareness of the discrepant response should result in a momentary pausing in behavior, heightened negative self-directed affect, and noting of stimuli present or in some way related to the discrepant response (e.g., where she was; where she was going; which purse she carried;
features of the man, such as his race). These consequences should work in concert to establish cues for control, or the building of associations between the prejudiced response (in this case, clutching the purse), the negative affect, and related stimuli (e.g., the man’s race). Theoretically, this process should take only milliseconds.

These initial consequences of becoming aware of a discrepant response should be critical for the self-regulation of prejudiced responses in the future. Specifically, the presence of cues for control in subsequent situations when a prejudiced response is possible should trigger the BIS, resulting in behavioral inhibition that allows one to engage in prospective reflection. In other words, the detection of a situation where a biased response may occur should interrupt ongoing responding and allow for a more careful consideration of how to respond. This enables one to inhibit a prejudiced response and generate a nonbiased response instead.

Theoretically, with practice, this process of self-regulation should result in the deautomatization of prejudiced responses and the consistent generation of less biased responses. That is, the capacity for control should increase with the accumulation of associations that trigger the biased responses that demand control. In the above example with Elaine, her previously established cues for control should trigger self-regulation so that she can interrupt the process of automatic bias and experience greater success responding in low-prejudiced ways. With additional practice, both upward and downward control can become routinized, making the self-regulation of prejudiced responses an automatic activity (see Lewis & Todd, 2007).

We turn now to a summary of findings from studies that have tested various aspects of the SRP model.

**Prejudice-Related Discrepancies**

A first step in investigating the SRP model involved determining the extent to which people are aware of their prejudice-related discrepancies, which is essential to cue development and subsequent self-regulation. A frequently used method for examining awareness of prejudice-related discrepancies involves administration of the Should–Would Discrepancy Questionnaire (e.g., Devine, Monteith, Zuwerink, & Elliot, 1991; Monteith, Devine, & Zuwerink, 1993; Monteith & Voils, 1998; see Monteith & Mark, 2005, for a review). Participants consider various situations in which biased responses are possible, and they make two types of ratings for each situation. First, they report the extent to which they would have the biased responses. Second, they report the extent to which they should have the biased response based on their personal standards for responding, defined as what they personally consider appropriate based on their beliefs. For example, White participants report the extent to which they would feel uncomfortable when shaking the hand of a Black person, and the corresponding should item asks about the extent to which they should feel uncomfortable doing so. Discrepancy scores are then generated by subtracting each should rating from the corresponding would rating and summing the resulting difference scores. This research has revealed that the vast majority of participants (approximately 80%) report positive discrepancies, suggesting that they are aware that they are prone to responding in ways that are more prejudiced than their personal standards suggest are appropriate.

However, are these self-reported discrepancies authentic, mapping on to actual discrepant responses? To address this question, Monteith and Voils (1998, Study 3) assessed Should–Would Discrepancies in an initial session, and in a separate session
examined how favorably participants evaluated jokes that played on stereotypes of Blacks. Reasoning that stereotypes were more likely to form a basis for responses to the jokes when cognitive resources were taxed (e.g., Pratto & Bargh, 1999), participants were put under low- or high-cognitive load while they evaluated the jokes. The results indicated that the low-prejudice participants who self-reported larger prejudice-related discrepancies experienced difficulty responding without bias under high-cognitive load. That is, these participants evaluated the jokes relatively favorably when they were cognitively taxed (i.e., a discrepant response), but generated less favorable evaluations under low load. The low-prejudice participants whose discrepancies were smaller (i.e., who reported that they would and should respond with little bias) generated unfavorable evaluations of the racial jokes, even when they were cognitively taxed. Thus, these results indicate that participants' reports of their degree of difficulty in controlling prejudiced responses, as indexed by their Should–Would Discrepancy scores, corresponded with their behavioral responses to the racial jokes.

Monteith, Mark, and Ashburn-Nardo (2010) used an interview procedure to assess Whites' proneness to discrepant responses toward Blacks in their everyday lives. Participants described experiences they had—if any—in which they felt, thought, or did something in relation to Blacks and later wished they had responded differently. Discrepancy experiences involving behaviors (e.g., avoiding a party because Blacks would be there), thoughts (e.g., thinking Black students do not study), or feelings (e.g., negative reaction to an interracial couple) were reported by most participants. Furthermore, their frequency was positively related both to Should–Would Discrepancy scores and to scores on Carver and White's (1994) BIS scale, which assesses the propensity to monitor and be sensitive to situations involving threats or punishment. Finally, reports of more discrepancy experiences in everyday life were obtained among people who were either internally or externally motivated to avoid prejudiced responses, compared to participants who were unmotivated to avoid prejudiced responses. In sum, this research provided particularly vivid and candid evidence of people’s awareness of their prejudice-related discrepancies, and awareness varied in meaningful ways with relevant individual difference variables.

Related Affect

Self-insight, however, does not automatically cure prejudice. At best it starts the individual wondering. And unless one questions the truth of his convictions, he certainly is unlikely to alter them. If he begins to suspect that they are not in conformity with facts, he may then enter a period of conflict. If the dissatisfaction is great enough, he may be driven to a reorganization of beliefs and attitudes.

Allport, 1954, pp. 328–329

Allport's (1954) seminal writing on the experience of prejudice with compunction suggests that people must be dissatisfied with their prejudice to be motivated to change. The SRP model likewise posits that awareness of one's discrepant responses must give rise to feelings of negative self-directed affect if such awareness is to be useful for ultimately learning to inhibit and change one's prejudiced patterns of responding. The affective consequences of discrepancy awareness have been examined in many studies (see Monteith & Mark, 2005, for a review) by having participants report their current
feelings immediately after they complete the Should–Would Discrepancy Questionnaire. To the extent that participants report prejudice-related discrepancies on the questionnaire, these inconsistencies should be primed and have affective consequences. Results have consistently indicated that low-prejudice individuals (and high IMS participants, see Plant & Devine, 1998) with larger discrepancy scores report greater general discomfort and also negative self-directed affect when they have just had their discrepancy proneness brought to mind. In contrast, as high-prejudice individuals’ proneness to discrepancies increases, they report more general discomfort but not feelings, suggesting that they are disappointed with themselves. This difference as a function of prejudice level is to be expected because low-prejudice individuals’ personal standards for responding are well-internalized and involve strong feelings of moral obligation, whereas this is less the case for high-prejudice individuals (e.g., Monteith et al., 1993; Monteith & Walters, 1998).

Recall that Monteith et al.’s (2010) interview method involved participants recounting discrepancy experiences they had in their everyday lives. As people described the situations involving their discrepancies, 64% talked about negative self-directed affect they experienced over their discrepant response (e.g., saying they felt guilty, morally wrong, ashamed, dirty inside, upset with myself, hypocritical) and 55% referenced general discomfort (e.g., saying they felt bothered, awkward).

The affective consequences of awareness of prejudice-related discrepancies have also been investigated experimentally by leading participants to believe they have engaged in prejudiced responses. For example, Monteith (1993) led heterosexual participants who evaluated the credentials of a supposed law school applicant to believe that their negative evaluation of him was based on his being gay. In another series of investigations (Monteith et al., 2002), low-prejudice White participants were given fixed physiological feedback suggesting that they had negative reactions when viewing racial pictures (e.g., an interracial couple). Fehr and Sassenberg (2010) provided German participants with IAT feedback indicating that they had negative attitudes toward Arabs. Such experimental research has consistently shown that low-prejudice participants and participants with high internal motivation to respond without prejudice report significantly more negative self-directed affect when they believe they are having prejudiced responses that conflict with their personal standards than when they do not.

In an intriguing investigation, Amodio, Devine, and Harmon-Jones (2007) tested a dynamic conceptualization of discrepancy-associated guilt, hypothesizing that it initially functions as a negative reinforcement cue and reduces approach motivation, but then it transforms into approach-motivated behavior when one has the opportunity for reparation. Whereas the initial reduced approach motivation helps one to interrupt ongoing behavior and process the transgression to learn from mistakes, the transformation of guilt into approach responses facilitates more personally acceptable responding in the future. Because frontal EEG has been validated as an index of motivation orientation in past research (Hagemann, Naumann, Thayer, & Bartussek, 2002), Amodio et al. reasoned that an EEG measure of frontal cortical asymmetry could capture changes in approach/withdrawal orientations linked with the two functions of guilt.

This conceptualization was tested among low-prejudice Whites who initially viewed faces of Whites, Asians, and Blacks while EEG recordings were made. Then they were given bogus feedback suggesting that they had moderately negative reactions when viewing faces of Blacks and more positive reactions to the other faces. A post-feedback assessment of guilt indicated that it was elevated relative to participants’ earlier baseline
levels. Furthermore, elevated guilt levels were associated with a reduction (relative to baseline) in left-sided frontal asymmetry, indicating a reduction in approach motivation. Participants were then asked to evaluate magazine article titles and indicated how much they would be interested in reading each article. Three titles were relevant to prejudice reduction (e.g., "10 ways to reduce prejudice in everyday life"). When this opportunity for approaching a prejudice-reducing activity was introduced, elevated guilt levels were associated with greater self-reported interest in the activity and also with increased left-frontal asymmetry, indicating an increase in approach motivation.

**Behavioral Inhibition and Retrospective Reflection**

The momentary interruption of ongoing behavior posited to occur by the SPR model upon the detection of discrepant responses has been indexed with reaction times. For instance, in Monteith et al.'s (2002) fixed physiological feedback experiments, pictures of Blacks and neutral pictures were presented to participants on a computer screen one at a time, and each one was followed by a graph that ostensibly depicted participants' supposed level of negative arousal. Participants were instructed to press the spacebar after seeing each graph to move on to the next picture, and the time taken to do so was measured (in milliseconds) as an indicator of behavioral inhibition. As expected, participants who received feedback suggesting that they were having negative reactions to the racial pictures took longer to press the spacebar than participants who received the same type of feedback but for the neutral pictures.

In contrast to behavioral inhibition, which is an interruption of ongoing behavior, retrospective reflection involves paying attention to features of the discrepancy situation and processing information that may help to predict similar discrepant responses in the future. This aspect of the SPR model has been tested primarily in thought-listing tasks following the induction of a prejudice-related discrepancy. For example, in the fixed physiological feedback studies noted above (Monteith et al., 2002), participants listed their thoughts about the experiment at the end of the study. Participants who believed they had negative reactions to pictures, including Blacks, were more preoccupied with their reactions to the pictures than participants who believed they had negative reactions to the neutral pictures. Similar findings were obtained by Monteith (1993, Study 1), in that low-prejudice participants were especially preoccupied with their negative evaluations of a law school applicant when they believed that their evaluations reflected a heterosexist bias.

**Conflict Monitoring and Detection: Neuroscientific Evidence**

The SRP model maintains that people will engage in conflict monitoring to detect potentially biased responses after establishing cues for control, a process that appears to be mediated by activity of the left prefrontal cortex (LPFC) and the ACC (Botvinick et al., 2001) and thus most appropriately examined with neural measures. For instance, Cunningham, Johnson, Raye, Gatenby, Gore, and Banaji (2004) presented Black and White faces to low-prejudice White participants during event-related functional magnetic resonance imaging (fMRI). When Black faces were presented for 525 milliseconds (but not at shorter intervals), heightened activity in the LPFC was observed, suggesting monitoring activity.

Using a different paradigm, Amodio, Harmon-Jones, Devine, Curtin, Hartley, and Covert (2004) monitored electroencephalographic activity while White, low-prejudiced
participants completed the weapons identification task (Payne, 2001)—a measure of implicit racial bias. The researchers were interested in the error-related negativity (ERN) wave, which appears to originate from activity in the ACC (Dehaene, Posner, & Tucker, 1994). The ERN is a component of the event-related potential that is sensitive to the conflict-detection process, and especially to conflicts that can result in failures to implement control (Gehring & Fencsik, 2001; van Veen & Carter, 2002). Amodio et al. reasoned that ERNs would be larger when participants made racial bias errors on the weapons task (i.e., when they mistakenly “shot” Blacks holding tools) relative to when they made errors on comparison trials (i.e., on White-tool trials). The results confirmed this prediction. In addition, process dissociation procedures (Jacoby, 1991) were used to separate the extent to which participants’ performance on the weapons task was due to automatic versus controlled processing. The researchers found that the greater participants’ ERNs to race-biased responses, the greater their control while completing the task (see also Amodio, Devine, & Harmon-Jones, 2008; Bartholow, Dickter, & Sestir, 2006; Legault & Inzlicht, 2013). Thus, the findings not only point to discrepancy-induced conflict activity but also to a link between this activity and greater subsequent controlled regulation of race bias.

Evidence of conflict monitoring has also been observed among people who are motivated to avoid prejudiced responses for external reasons. Amodio, Kubota, Harmon-Jones, and Devine (2006) reasoned that whereas the conflict-monitoring processes involved in race-bias regulation among high IMS individuals is related to activity in the dorsal regions of the ACC, externally motivated individuals should show a different pattern of neural activity, especially when there are strong external cues for controlling prejudiced responses. The rostral subregions of the ACC appear to be responsive to perception of response errors and the processing of external goal contingencies (e.g., Garavan, Ross, Kaufman, & Stein, 2003), and so Amodio et al. (2006) anticipated that this region would be related to stereotype inhibition among high EMS participants when they were responding in public. This was indeed the case.

As Amodio et al. (2008) argued, neurocognitive research provides a “non-humuncular” explanation for how the need for control can be signaled. Furthermore, these detection and control processes are able to tune perception (Amodio, 2010) and operate below conscious awareness (Amodio et al., 2004), making them critical to efficient and effective self-regulation.

**Prospective Reflection and Prejudice Regulation in the Presence of Cues for Control**

When the monitoring process detects the presence of cues for control in contexts where discrepant responses are possible, the SRP model holds that the cues trigger an interruption of ongoing responding (prospective reflection) and the generation of non-prejudiced responses. In one study relevant to these hypotheses (Monteith, 1993, Study 2), low-prejudice heterosexual participants received feedback on a (bogus) test of subtle prejudice indicating that they were prone to subtle prejudice in relation to gays. This constituted the initial prejudice-related discrepancy experience. The critical test of the prospective reflection and inhibition of prejudiced responses then occurred in a supposedly separate second study about humor, where 2 of 12 jokes that participants evaluated played on stereotypes of gays. The results indicated that participants given discrepancy feedback took longer (in milliseconds) to evaluate these anti-gay jokes and evaluated them less favorably relative to participants in a control condition who had not been given feedback that they were prone to subtle prejudice.
In another study that tested the cues for control idea in a different way (Monteith et al., 2002), low-prejudice participants took a racial IAT that involved the dual categorization of traditionally Black and White names, and pleasant and unpleasant words. They were given performance feedback suggesting racial bias, which elicited negative self-directed affect. Later in the study, as part of a supposedly unrelated task, participants were presented with words one at a time on the computer and were asked to indicate their first reaction to each word—was it something they liked or disliked? Some words were traditionally Black names from the IAT. Results indicated that participants paused longer for these names to the extent that their performance on the IAT had elicited negative self-directed affect. Furthermore, the more guilt experienced in relation to IAT performance, the more positively participants evaluated the traditionally Black names in the like/dislike task. These results suggest that the more participants felt guilty about their initial discrepancy, the more the race-related stimuli from that task triggered an interruption of responding and unbiased evaluations.

**Application to High-Prejudice Individuals**

Although the SRP model was developed with people who are personally motivated to avoid prejudiced responses in mind, it can also apply to regulation among individuals who are more prejudiced. Even if one’s conscious attitudes are negative toward a particular group, pressures to respond with less prejudice and punishments for failures to do so (e.g., social rejection) likely prompt the development and operation of cues for control among high-prejudice individuals. Thus, external motivations, such as motivations to avoid social sanction or to gain social approval (Legault et al., 2007) can lead to prejudice regulation. The regulation is driven more by a desire to avoid discomfort and admonishment than by negative self-directed affect (e.g., Monteith et al., 1993; Plant & Devine, 1998). Also, cues for control are likely to involve public settings where prejudiced responses are possible (Plant & Devine, 1998) or people who may sanction one’s biases. This was evident in Monteith et al.’s research (2010) in which White participants were led through a guided interview designed to gather information about the development and operation of cues for control. For example, one high-prejudice participant reported:

My roommate’s Black and sometimes when we’re watching shows they kinda like make the Blacks look trashy, you know like on Jerry Springer . . . I was laughing at it but he wasn’t really and it kind of automatically made me feel like I had done something wrong so I felt bad . . . I didn’t want him to think, “Well he looks like some kind of racist.”

When asked whether this experience affected the participant in the future, he responded:

If something on TV comes up that’s like shady you know it’s like I think about it . . . you know I think about it to make sure that it doesn’t happen again in case he actually was mad about it. I wouldn’t laugh out loud if I thought maybe it would be offensive to someone else. I’m just a little more careful now.

In sum, both high- and low-prejudice individuals seem to put the “brakes” on their prejudices (Allport, 1954) through the development and operation of cues for control, although this process is driven more by external than internal motivations for more prejudiced people.
SELF-REGULATION THROUGH SUPPRESSION

An alternative to the SRP strategy of self-regulating prejudice that might seem more straightforward and efficient involves simply banishing stereotypic and biased thoughts from the mind. This is like the dieter who says, "I'm simply not going to think about food." Stereotype (or prejudice) suppression involves quite simply trying not to have stereotypic (or prejudiced) thoughts and instead focusing on "distracter" thoughts.

Unfortunately, the active effort involved in attempting to banish stereotypic thoughts from the mind has often been shown to backfire and result in a rebound effect. For example, in the classic demonstration of this effect, Macrae, Bodenhausen, Milne, and Jetten (1994, Study 1) asked participants to spend five minutes writing about a typical "day in the life" of a skinhead shown in a photograph. Half the participants were instructed to avoid stereotypic thoughts while writing their passage, and others were given no special instructions. Participants then wrote another passage, again about a skinhead, but this time none of them were instructed to suppress stereotypes. Results indicated that participants who had initially suppressed stereotypes showed a rebound effect on the second passage they wrote, such that they used stereotypes even more than participants who never received suppression instructions. This rebound effect was also replicated with behavioral (Study 2) and stereotype accessibility (Study 3) measures.

The theoretical explanation for the stereotype rebound effect relates to Wegner's (1994) model of mental control. This model posits that, while individuals engage in a controlled operating process in an attempt to regulate their thoughts by identifying appropriate distractor thoughts, an ironic monitoring process continually searches for evidence of these unwanted thoughts. Unfortunately, this has the effect of priming the unwanted thoughts (Macrae et al., 1994), and when the conscious monitoring process is taxed or relaxed, the suppressed thought will return "with a vengeance."

Many studies have further demonstrated the paradoxical effects of stereotype suppression. They occur not only with blatant experimenter instructions to avoid stereotyping, but also when more subtle situational cues prompt suppression attempts (Macrae, Bodenhausen, & Milne, 1998). They are manifested in superior memory for stereotypical behaviors and impaired memory for nonstereotypic individuating information (Macrae, Bodenhausen, Milne, & Wheeler, 1996; Sherman, Stroessner, Loftus, & Glenn, 1997). Stereotype suppression effects even show up when suppressors perform tasks that are unrelated to the target group, such as when participants who suppressed stereotypes about sportsmen (associated with poor math performance) subsequently performed worse on a calculus task than non-suppressors, and people who suppressed stereotypes of the elderly subsequently walked slower than non-suppressors (Follenfant & Ric, 2010). Also, because effortful stereotype suppression depletes regulatory resources, it can increase stereotyping in general, and not just in relation to the group about which stereotypes were initially suppressed (Gordijn, Hindriks, Koomen, Dijkstra, & Van Knippenberg, 2004).

This body of research suggests that self-regulating one's stereotypic thoughts through suppression will not only be unsuccessful but will also have counterproductive outcomes. However, researchers have also identified important boundary conditions to the stereotype rebound effect (Monteith, Sherman, & Devine, 1998), finding it is a consistently successful strategy for some people, and for others it can be successful under some conditions. Using the same "day in the life" paradigm as Macrae et al. (1994), Monteith, Spicer,
and Tooman (1998) found that low-prejudice individuals were less prone to the rebound effect. Gordijn et al. (2004) similarly found that individuals who are high in internal motivation to suppress stereotypes did not reveal the typical rebound effect. They further showed that these individuals did not show a depletion of regulatory resources following stereotype suppression, whereas people low in suppression motivation did. Therefore, the act of suppression does not appear to be as taxing for internally motivated people, perhaps because egalitarian thoughts provide ready replacements (see Monteith, Spicer et al., 1998).

Even high-prejudice people do not show the stereotype rebound effect if social norms call for continual avoidance of the use of stereotypes after the initial suppression period and if cognitive resources are sufficient for continued suppression (Monteith, Spicer & Tooman, 1998; see also Wyer, Sherman, Strossner, 2000). Furthermore, if self-control is strengthened for high-prejudice individuals even in a domain unrelated to stereotype use, stereotype rebound can be avoided. Gailliot, Plant, Butz, and Baumeister (2007) found that participants who were low in motivation to control prejudice had to exert quite a lot of self-control to suppress stereotypes initially, as evidenced in impairment in executive functioning following stereotype suppression (measured by Stroop and anagram task performance). However, after they had completed two weeks of self-regulation exercises in a stereotype-unrelated domain (e.g., refraining from cursing), the act of stereotype suppression no longer depleted regulatory resources.

In sum, stereotype suppression may sometimes prove successful and not result in rebound. Nonetheless, we believe that this strategy has a potentially important limitation beyond the possibility of producing rebound effects. Because it does not create a positive goal that one works toward but rather places a focus on avoidance of unwanted thoughts and outcomes, this strategy alone is not likely to prove effective in producing long-term changes to the stereotypic and evaluative underpinnings of prejudiced responses even among low-prejudice individuals. Furthermore, research has indicated that encouraging people to comply with pro-egalitarian pressures when their personal preference is for prejudiced responding can lead to anger and backlash (Plant & Devine, 2001).

IMPLEMENTATION INTENTIONS AND SELF-REGULATION

People may also self-regulate their prejudice through a more direct path to the distal goal of egalitarian responding that is not so dependent on monitoring, inhibition, or suppression processes. This method involves establishing implementation intentions (Fujita, 2011; Gollwitzer, 1999; Gollwitzer & Sheeran, 2006), which are if–then plans that spell out contingencies between situations and responses (e.g., “If situation Y is encountered, then I will initiate egalitarian response X”). By creating a mental link between a cue or situation (e.g., “If I see a Black person in a store”) and a goal-directed response (e.g., “then I will assume that the person is shopping and not working at the store”), one becomes committed to performing the response. In addition, mental representations of specified cues become more accessible when implementation intentions are formed (e.g., Aarts, Dijksterhuis, & Midden, 1999) so that the intended response can be executed immediately and reflexively (Gollwitzer & Brandstätter, 1997) and with little conscious intent (Bayer, Achtziger, Gollwitzer, & Moskowitz, 2009). Implementation intentions also help to shield people from the intrusion of unwanted thoughts (Achtziger, Gollwitzer, & Sheeran, 2008), and appear to be pursued tenaciously while conserving self-regulatory
resources (Martijn, Alberts, Sheeran, Peters, Mikolajczak, & de Vries, 2008). Support for this method of self-control is abundant across many domains, including health screening, completing academic assignments, dieting behavior, and environmental behaviors such as recycling (see Gollwitzer & Sheeran, 2006). It also has promise for the self-regulation of prejudiced responses.

Thus far, the use of implementation intentions in regulating intergroup bias has been examined in the context of implicit stereotyping tasks. Stewart and Payne (2008) had participants form the intention to think of a certain counterstereotypical thought related to Blacks. Specifically, in two experiments participants were instructed to think “safe” when completing the weapons-identification task (Payne, 2001), which measures associations between weapons versus neutral objects, and Blacks versus Whites. In a third experiment, participants were instructed to think “good” when they saw a Black face while completing a race-based IAT. Each experiment revealed reduced evidence of implicit bias when participants were given the implementation intention, and process dissociation analyses showed that the implementation intention reduced automatic stereotype bias (rather than increasing controlled processing).

Mendoza, Gollwitzer, and Amodio (2010) examined the effectiveness of implementation intentions in the context of performance on the Shooter Task (Correll, Park, Judd, & Wittenbrink, 2002). This task involves quickly shooting armed targets shown on the computer screen and not shooting unarmed targets. The targets are Black on some trials and White on others. The typical bias is that people mistakenly shoot unarmed Black targets significantly more than unarmed White targets. Participants in a first experiment were given the distraction-inhibiting implementation intention: “If I see a person, then I will ignore his race!” Relative to a control condition that was not given this implementation intention, participants in the intention condition were significantly less likely to shoot the unarmed Black targets; indeed, they made just as few errors on unarmed Black target trials as on unarmed White target trials. Process dissociation analyses indicated that automatic processing decreased and controlled processing increased for Black targets in the implementation intention strategy condition, relative to the control condition. In a second experiment, a very subtle difference was introduced between a simple-goal strategy condition and an implementation strategy condition. Specifically, in the simple-goal condition, participants were instructed to use the strategy “I will always shoot a person I see with a gun!” and “I will never shoot a person I see with an object!” The implementation intention conditions were given these strategies in if–then terms: “If I see a person with a gun, then I will shoot!” and “If I see a person with an object, then I will not shoot!” Results indicated that the implementation intention strategy increased performance accuracy overall, thus reducing the influence of implicit stereotypes on responses. Process dissociation analysis showed that controlled processing was enhanced by the implementation intention.

The use of implementation intentions to regulate prejudice in everyday life may be somewhat tricky for a couple of reasons. First, the use of blanket implementation intentions such as “If I see a person of color, then I will ignore his/her race!” may frequently be inadvisable, because people’s group membership is after all frequently an important part of their identity, and also a colorblind approach is known to produce a number of negative consequences (Apfelbaum, Sommers, & Norton, 2008). Second, coming up with if–then behavioral plans with cross-situational applicability may be a challenge. Nonetheless, use of implementation intentions along with other regulatory strategies may have
more generalized value. As Mendoza et al. (2010) pointed out, once one has identified cues for control through self-regulation, these cues can be linked to the intention to respond carefully—to slow down and take in details of the situation so that an appropriate egalitarian response can be initiated. Thus, people may form the implementation intention, “If I see a person of color, I will think before acting on stereotypes.”

We believe that implementation intentions can also be fruitfully used in concert with a strategy for implicit bias reduction called counterstereotyping (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000). Counterstereotyping has been studied as a means of conditioning associations that run contrary to stereotypes by having participants repeatedly pair the outgroup (e.g., women) with counterstereotypes (e.g., leader) across hundreds of trials. It has been shown to reduce stereotype accessibility, resulting in less biased performance on implicit measures. How might this strategy be implemented in everyday life, rather than through intensive computer-mediating conditioning? Recently in our lab, we have taught people to generate counterstereotypes in their everyday lives by using the implementation intention “If I see a Black person, then I will think of a counterstereotype.” Participants practiced countereotyping across two weeks, and content analysis of participants’ descriptions of their use of the strategy suggested that they applied it successfully in everyday contexts. Furthermore, use of the countereotyping strategy was associated with reduced automatic bias on implicit measures of bias a full two months after the initial two-week practice period.

**SELF-CONTROL DEPLETION AS A THREAT TO THE SELF-REGULATION OF PREJUDICE**

An influential perspective that is related to people’s ability to engage in self-regulation is Baumeister and colleagues’ strength model, which maintains that self-control is a limited resource resembling a muscle (Muraven & Baumeister, 2000; for a meta-analysis, see Hagger, Wood, Stiff, & Chatzisarantis, 2010). Although self-control strength is necessary for the executive functioning that is involved in self-regulation, self-control resources can be depleted by engaging in self-regulation, or having multiple regulatory demands or stressors. Just as when a muscle becomes fatigued so that less physical effort can be exerted, regulatory activity likewise depletes people’s ability to engage in regulatory control. The depletion of regulatory resources has been linked with drops in blood glucose levels, which are apparently important for providing “brain energy” for engaging in self-regulation (Galliot, Baumeister, deWall, Maner, Plant, Tice, Brewer, & Schmeichel, 2007). Restoring self-control resources by giving people sucrose drinks thus can have positive regulatory effects, including reduced stereotyping (Galliot, Peruche, Plant, & Baumeister, 2009). More direct evidence of neural depletion was obtained by Inzlicht and Gutsell (2007), who showed that error-related negativity (ERN), which is a waveform associated with anterior cingulate activity, yields weaker signals following the exertion of self-regulatory efforts. Thus, it appears that the brain can only handle so much regulation, and then self-control may fail.

Does the dependence of the regulatory system on sufficient control strength for its operation threaten the likelihood that people will successfully self-regulate their prejudice? This is clearly possible, although there are two important things to keep in mind. First, certain self-regulatory actions related to prejudice control may be more likely to lead to depletion than others. The pursuit of avoidance goals in particular is likely to lead to depletion across time (Oertig, Schuler, Schnelle, Bandstatter, Roskes, &
Elliot, 2013). Thus, people focused solely on monitoring and preventing bias from occurring (e.g., as with stereotype suppression), without also focusing on the goal of responding in egalitarian ways, may be most prone to experiencing self-regulation failure. When people are “action-oriented” (i.e., high-level goals and intentions exert a “top down” influence), demanding situations can actually improve self-regulation by mobilizing controlled processes (Koole, Jostmann, & Baumann, 2012).

Second, there are a variety of ways to preserve and replenish the self-control reservoir. Understanding what helps to maintain and increase self-control strength can help us to determine how people can become more effective self-regulators. As with a muscle, self-control can be strengthened by exercising it across time (Muraven, Baumeiser, & Tice, 1999). For instance, Gailliot et al. (2007) found that two weeks of self-regulatory exercises that were unrelated to prejudice (e.g., using one's nondominant hand or refraining from cursing) prevented the self-regulatory depletion effect typically observed following stereotype suppression. Control strength can also be increased by self-affirming one's core values (Schmeichel & Vohs, 2009). Specifically, thinking about one's core values promotes abstract or high-level mental construal, which enhances self-control. Thus, reminding one's self of one's core egalitarian values when engaged in the self-regulation of prejudice should protect against depletion. There is also evidence that people can overcome the effects of regulatory depletion if they believe that persistence will improve one's abilities (Muraven & Slessareva, 2003), believe that the capacity for self-control is not limited (Job, Dweck, & Walton, 2010), expect that persistence will lead to improvement (Martijn, Tenbult, Merckelbach, Dreesens, & de Vries, 2002), or monitor their performance in light of their standards (Wan & Sternthal, 2008). We believe that this body of research suggests that depletion effects can often be overcome during the self-regulation of prejudice, particularly among people who are personally committed to egalitarianism.

SELF-REGULATION IN INTERGROUP INTERACTIONS

An important context in which the depleting effect of regulation has been studied is interracial interactions. Such interactions can be stressful and anxiety-provoking (e.g., Plant, 2004), especially when attempting to self-regulate one's responses to avoid responding in prejudiced ways or being perceived as biased. Blascovich, Mendes, Hunter, Lickel, and Kawai-Bell (2001; see also Amodio, 2009) found that nonstigmatized individuals experienced greater levels of threat (as measured via self-report, behaviorally, and physiologically) during an intergroup interaction than during an intragroup interaction. Threat responses were especially evident to the extent that the participants had little prior intergroup contact experience.

Intergroup interactions can also take a toll on executive functioning because they deplete regulatory resources. Richeson and Shelton (2003) provided an elegant test of this idea by having White participants interact with a White or a Black experimenter after they had just completed the race IAT. During the interracial interaction, participants who scored higher on race IAT moved their body less, looked around the room less, and moved their hands less than participants with lower implicit bias scores if they were in the Black experimenter condition. This suggests that participants were attempting to regulate and control their behavior during the interracial interaction to the extent that they had implicit bias. After the interaction, participants completed a Stroop task as a measure of cognitive interference. The results indicated that the greater participants' bias
on the race IAT, the greater interference they experienced on the Stroop task following their interaction with the Black experimenter. In contrast, bias on the IAT was unrelated to cognitive interference in the White experimenter condition. In sum, participants who were more prone to implicit race bias appeared to regulate their behavior more when interacting with a Black than a White experimenter, and they also showed reduced executive functioning as indexed by interference on the Stroop task.

In a related investigation, Richeson, Baird, Gordon, Heatherton, Wyland, Trawalter, and Shelton (2003) conducted a two-session study with White participants. The first session involved a replication of Richeson and Shelton (2003) findings as described above. In a second session with the same participants, functional magnetic resonance imaging (fMRI) data of the right dorsolateral prefrontal cortex (DLPFC) and the ACC were collected while participants viewed pictures of unfamiliar Black and White faces. Richeson et al. (2003) reasoned that because these brain regions support executive control processes such as response monitoring, inhibition, and conflict detection (e.g., Cohen, Botvinick, & Carter, 2000), they should be activated upon viewing Black faces to the extent that participants were more prone to implicit racial bias. Results indicated that this was indeed the case. Furthermore, analysis of the relation between data collected in Sessions 1 and 2 indicated that neural activity was related to interference on the Stroop task. The critical test in this research then involved an examination of whether the extent of activity in the brain regions thought to be critical to executive control mediated the relation between implicit racial bias and Stroop interference. The results supported the mediated test (in particular, for activity in the DLPFC). These findings suggest that the interracial interaction was cognitively depleting for participants who were prone to implicit racial biases because it taxed limited executive control resources.

This resource depletion account has also been supported in investigations that manipulated the self-regulatory demands of interracial interactions. In Richeson and Trawalter’s (2005) research, White participants’ concerns about being prejudiced were manipulated prior to a dyadic interaction by having participants complete the racial IAT and telling them that “most people are more prejudiced than they think they are.” Other participants were given prejudice-unrelated feedback after completing the IAT. Later, participants in the prejudice concern condition showed more impairment of Stroop performance if they had interacted with a Black confederate than if they had interacted with a White confederate. In contrast, less Stroop interference was evident as a function of race of interaction partner among participants whose prejudice concerns had not been heightened.

Despite the cognitive burden that appears to result from the regulation of implicit prejudice, efforts at self-regulation can ultimately have positive effects on how individuals are perceived by outgroup members. Recall that Richeson and Shelton (2003) found that bodily control (less eye, hand, and body movement) during an interracial interaction increased as proneness to implicit racial bias increased. Another study (Shelton, 2003) similarly revealed that White participants who were instructed to “try not to appear prejudiced” during an interaction with a Black participant fidgeted less than participants who were not given this instruction. In addition, Black participants reported more favorable evaluations of White participants who were asked to try to control their prejudice. Given these findings, Shelton, Richeson, Salvatore, and Trawalter (2005) reasoned that the behavioral control exerted by Whites who are prone to implicit racial biases might lead Black individuals to perceive them as quite engaged during interracial
interactions, which may lead to greater liking. To test this hypothesis, White participants first completed the racial IAT. Then they interacted either with a White or Black partner, and partners subsequently reported how engaged the White participants were in the interaction and how favorably they felt toward these individuals. The results revealed that Whites who were more prone to implicit racial bias were perceived more favorably by Black (but not White) interaction partners. Furthermore, this relation was mediated by the extent to which the Black interaction partners perceived the White participants to be engaged in the interaction.

Vorauer and Turpie (2004) report another line of research suggesting that individuals who are higher in explicit prejudice are likely to engage in positive interracial behaviors out of concerns about appearing prejudiced. White Canadian participants prepared a videotaped message that supposedly would be shown to another participant who was identified either as a White or a First Nations Canadian student. In addition to varying the ethnicity of participants’ supposed partner (i.e., the person who would view the videotape), the researchers examined the effects of participants’ evaluative concerns. For example, high evaluative concern was created by telling participants that they should “watch themselves” while recording their message. Vorauer and Turpie (2004) reasoned that these concerns would lead participants to exhibit fewer intimacy-building behaviors (e.g., self-disclosure) and behaviors that convey positive feelings and foster closeness. The findings indicated that participants who were higher in explicit prejudice “shined” in their behaviors; in other words, they were more likely to engage in intimacy-building behaviors to the extent that they were concerned about how they would be evaluated by their supposed partner. In contrast, people who were lower in explicit prejudice “choked,” such that they were less likely to engage in the intimacy-building behaviors to the extent that they were concerned about their partner’s evaluations.

In sum, the literature examining self-regulation during intergroup interactions suggests that it can have beneficial outcomes for the quality of the interaction, even if it has a depleting effect on the regulator. In our view, a little cognitive depletion is a small price to pay when people are practicing bias regulation during interracial interactions. Furthermore, the extent of depletion is linked to the strategies people use for self-regulation. As we have summarized herein, sometimes people carefully monitor their behavior and vigilantly attempt to prevent bias from entering into their responses. Whereas such a prevention focus is cognitively depleting, people who are focused instead on having a “positive intercultural exchange” (i.e., promotion focus) are not as cognitively taxed by interracial interactions (Trawalter & Richeson, 2006). As a final caveat, we suspect that depletion is a short-term consequence that will dissipate as people become more experienced at interracial interactions.

**USING INTERPERSONAL CONFRONTATION TO ENCOURAGE SELF-REGULATION**

A final topic we wish to address briefly concerns how people can take an active role in promoting the self-regulation of prejudiced responses among others. If our eyes and ears are open, we will observe numerous instances of bias in our daily lives either directly (i.e., in our own interactions) or indirectly (e.g., in the media). Rather than ignore your mother’s comment about “going to that side of town,” your coworker’s inclination to interview only male job applicants, or your friend’s avoidance of gays, you can speak up
and confront the other’s bias. For example, you might question the perpetrator, express surprise, communicate that you are taken aback, or explain the problem with a particular comment. Standing up for equality through interpersonal confrontation is something that people hesitate to do (Kawakami, Dunn, Karmali, & Dovidio, 2009; Swim & Hyers, 1999), especially when the potential costs of doing so are salient (Shelton & Stewart, 2004). However, interpersonal consequences of confrontation (e.g., liking) tend to be much more positive than people’s forecasts suggest they will be (Mallet & Wagner, 2011). Also, people can avoid feeling regretful and guilty about failing to confront (Shelton, Richeson, Salvatore, & Hill, 2006). Most importantly, confrontation is effective at getting others to self-regulate their prejudice (Czopp, Monteith, & Mark, 2006).

For instance, Czopp et al. (2006) examined people’s affective reactions to confrontation, and also how future behavior and self-reported explicit bias were affected. The task used to elicit responses that could be confronted involved presenting participants with a series of photos and accompanying descriptive statements. Participants then provided an attribution-like response for each photo-statement pair. On three trials, photo-statement pairs were used that were likely to elicit stereotypic responses (e.g., “This person relies on the government for money” rather than a non-stereotypic alternative such as “postal worker”). Afterwards, half of the participants were confronted by a confederate who pointed out how the participants’ responses were stereotypical, whereas the other participants were not confronted. Results showed that the confrontation caused heightened feelings of guilt and self-disappointment. In addition, participants who were confronted gave significantly fewer stereotypic responses on a similar picture-statement task at the end of the experiment, even though they believed these responses would be entirely private. Finally, people who were confronted reported significantly less explicit prejudice than those who were not confronted. This research demonstrates the powerful role that confrontations can play in promoting the self-regulation of prejudice.

Even if people are merely witnesses to a confrontation that does not involve them directly, the confrontation can encourage regulatory outcomes (Rasinsky & Czopp, 2010). Confrontations can be so powerful for a couple of reasons. First, they can increase awareness of subtle biases that might otherwise go undetected. Second, they establish or make salient egalitarian or nonprejudiced social norms. By pointing out prejudice, confrontations emphasize principles of fairness or equality. Boosting the strength of these norms and standards may be especially important for motivating people without strong egalitarian principles and values to self-regulate their prejudice.

People who are not targets of a particular form of bias may believe that it is not their business to confront. However, research indicates that these individuals may be especially effective confronters because they are not targets of the prejudice themselves, whereas members of stereotyped groups tend to be dismissed as complaining or overreacting (Czopp & Monteith, 2003; Gulkem, Mark, & Monteith, 2013; Rasinsky & Czopp, 2010). Precisely those individuals who think they should step aside should actually step up and use confrontation as a tool for encouraging self-regulation.

**CONCLUSION**

When Bargh reviewed the literature on stereotyping and the possibility of controlling automaticallyactivated stereotypes in 1999, his ultimate conclusion was not optimistic.


