Reducing the Expression of Implicit Stereotypes: Reflexive Control Through Implementation Intentions
Saaid A. Mendoza, Peter M. Gollwitzer and David M. Amodio
Pers Soc Psychol Bull 2010; 36: 512
DOI: 10.1177/0146167210362789

The online version of this article can be found at:
http://psp.sagepub.com/cgi/content/abstract/36/4/512

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
Society for Personality and Social Psychology

Additional services and information for Personality and Social Psychology Bulletin can be found at:

Email Alerts: http://psp.sagepub.com/cgi/alerts
Subscriptions: http://psp.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav
Citations http://psp.sagepub.com/cgi/content/refs/36/4/512
Reducing the Expression of Implicit Stereotypes: Reflexive Control Through Implementation Intentions

Saaid A. Mendoza, Peter M. Gollwitzer, and David M. Amodio

Abstract
The authors tested the effectiveness of implementation intentions as a strategy for limiting the behavioral expression of implicit stereotypes. Implementation intentions are if-then plans that link an intended response to an anticipated situational cue, thereby enabling reflexive control. In Study 1, participants used a distraction-inhibiting implementation intention designed to engage control over the perception of goal-irrelevant stimuli (e.g., race). In Study 2, participants used a response-facilitating implementation intention designed to promote goal-directed action. Across studies, implementation intentions improved accuracy, thereby limiting the behavioral expression of implicit stereotypes. Furthermore, process dissociation analyses indicated that the distraction-inhibiting implementation intention increased controlled processing while reducing automatic stereotype activation, whereas the response-facilitating implementation intention increased only controlled processing. Implications for goal strategy approaches to reducing prejudice are discussed.

Keywords
stereotypes, control, implementation intentions, goals, process dissociation

Received March 6, 2009; revision accepted August 25, 2009

Although Americans’ attitudes toward racial outgroups have improved in the past half century (Schuman, Steeh, Bobo, & Krysan, 1997), implicit stereotypes continue to influence people’s thoughts and behavior despite their egalitarian intentions (Bargh, Chen, & Burrows, 1996; Devine, 1989). Several influential theorists ascribe implicit stereotyping effects to the general process through which the mind organizes information: If humans rely on categories for the processing of complex social information, the use of stereotypes may be unavoidable (Allport, 1954; Taylor, 1981). Bargh (1999) noted, provocatively, that because effort and deliberation are required to overcome the use of automatic stereotypes, their effects are virtually inevitable. However, in the decade since Bargh’s chapter, advances in models of controlled processing suggest that with the aid of action plans, control can be engaged and implemented with little deliberative effort (Amodio et al., 2004; Lengfelder & Gollwitzer, 2001), and therefore, the activation and use of stereotypes may not be as inevitable as presumed. The present research integrated theories of prejudice and goal pursuit to address how implementation intentions can be used to spontaneously engage goal-directed responses and thereby limit the influence of automatic racial stereotypes on behavior.

Approaches to Reducing Implicit Stereotyping Effects
Expressions of racial bias are believed to comprise both automatic and controlled components (Devine, 1989), and either component may be targeted by strategies that aim to reduce the behavioral expression of implicit stereotypes. To date, the majority of research on prejudice reduction has targeted the automatic component. This approach aims to alter the nature of the semantic and/or evaluative associations that compose participants’ basic mental representation of racial outgroups. Therefore, the goal is often to replace negative associations of Blacks with more positive associations to reduce net expressions of anti-Black bias (e.g., Dasgupta & Greenwald, 2001; Gawronski, Deutsch, Mbirkou, Seibt, & Strack, 2008; Kawakami, Phillips, Steele, & Dovidio, 2007; Olson & Fazio, 2006). The emphasis on changing underlying associations is

1New York University, New York, NY, USA
2University of Konstanz, Konstanz, Germany

Corresponding Author:
David M.Amodio, Department of Psychology and Center for Neural Science, New York University, 6 Washington Place, New York, NY 10003, USA
Email: david.amodio@nyu.edu
consistent with the traditional view that deliberative control is difficult to implement in spontaneous responses, such as those measured by reaction time assessments of implicit race bias. However, given recent research demonstrating that goal-driven responses can be engaged spontaneously with little deliberation (Amodio & Devine, 2010), it appears fruitful to also consider interventions that target the engagement of control, in addition to focusing on changing automatic associations.

Theories of control emphasize the importance of goals and motivations (e.g., Fazio, 1990; Monteith, 1993). In a general sense, control is characterized by the successful implementation of a goal-consistent response in the face of distracting or countervailing influences. For example, most interracial interactions occur in the context of a goal that is unrelated to race. Prejudice emerges within these interactions when one’s behavior toward an outgroup member is changed on the basis of stereotypes. If one is able to focus on the goals of the interaction (e.g., to solve a particular problem), then implicit stereotypes have less of a chance of emerging in behavior (Amodio & Mendoza, 2010). The same principle applies to reaction time measures of implicit stereotyping, such as Payne’s (2001) weapons identification task. In this task, the participant’s goal is to simply categorize stimuli that appear on the computer screen as guns or tools while ignoring the Black and White faces that momentarily precede each of these objects. Implicit stereotypes are expressed in behavior when performance on the main task (object categorization) is biased because of the influence of racial stereotypes, such as the stereotype that Black people are dangerous. Control is thus defined by accurate (i.e., goal-consistent) task performance, despite the presence of potentially biasing racial stimuli (Amodio, Devine, & Harmon-Jones, 2008; Payne, 2005). According to this theoretical analysis, the expression of implicit stereotyping effects on rapidly unfolding responses can be reduced with a strategy that increases goal-directed performance accuracy. We describe such a strategy in the next section.

**Engaging Reflexive Control Through Implementation Intentions**

Although the notion of nondeliberative control has received little attention in the intergroup bias literature, it has been studied extensively in research on action plans and goal pursuit. For example, following the action phases model of goal pursuit (Gollwitzer, 1990; Heckhausen & Gollwitzer, 1987), Lengfelder and Gollwitzer (2001) distinguished between more- and less-deliberative forms of action control. Reflective action control involves conscious reflection in the pursuit of a goal and thus corresponds closely to the deliberative process that is assumed to be precluded on implicit measures of racial bias. Reflexive action control, by contrast, refers to the automatic initiation of goal-directed behavior, which can operate implicitly and under conditions of cognitive load. Although much research in social psychology has focused on reflective forms of control (i.e., Wegener & Petty, 1997; Wilson & Brekke, 1994), recent research has demonstrated that neural processes associated with response control may be engaged without deliberation within a few hundred milliseconds of a control-eliciting cue (Amodio et al., 2004, 2008), corroborating the plausibility of reflexive control strategies.

With regard to the present concerns, strategies that promote reflexive action control in intergroup contexts may be especially effective for limiting the expression of implicit stereotypes in behavior. One such strategy is an implementation intention, which is designed to facilitate responses that promote goal attainment. Whereas simple goal intentions state only what one wants to achieve, implementation intentions are consciously formed *if-then* plans that indicate the specific cognitive or behavioral response that is to be made at a specific time and place (Gollwitzer, 1993, 1999). That is, goal intentions are typically structured as “I plan to do goal-directed behavior *x*,” without reference to a specific triggering cue or the specific response to follow. Because simple goal intentions often require additional planning and deliberation at the time of implementation, they are more easily derailed by extraneous factors, such as distractions or social stereotypes. By comparison, implementation intentions use if-then contingencies to specify a triggering cue and the intended action to follow, using the following format: “If situation *x* occurs, then I will do goal-directed behavior *y*.” By linking a specific goal-directed response to a situational cue through an if-then statement, the cue becomes cognitively accessible (Aarts, Dijksterhuis, & Midden, 1999). When the cue is then encountered, the planned response can be implemented reflexively without any further planning or deliberation (Brandstätter, Lengfelder, & Gollwitzer, 2001; Webb & Sheeran, 2007; for a meta-analysis, see Gollwitzer & Sheeran, 2006). In the context of intergroup responses, the *if* component could refer to the appearance of a racial cue that represents the potential for racial bias (Monteith, 1993). The *then* component might specify a planned response that limits the influence of the racial associations on one’s intended behavior. Hence, implementation intentions may be thought of as a strategy for transforming consciously held egalitarian goals into reflexively triggered actions that limit the behavioral expression of racial stereotypes.

**Implementation Intentions for Inhibiting Distractions and Facilitating Action**

Implementation intention research has focused on two general types of control—inhibitory control and action control—which may be targeted by two different types of implementation intentions. *Distraction-inhibiting* implementation intentions are designed to engage inhibitory control, such as to inhibit the initial perception of potentially distracting or biasing
stimuli. For example, Gollwitzer and Schaal (1998) examined the effects of a distraction-inhibiting implementation intention on a gender stereotyping task. In the task, participants completed a stereotype Stroop task, in which they were to name the color of gender stereotypic and nonstereotypic words. Prior to each trial, a prime word was presented that consisted of a female name or a nonsense word. In a control condition, the female name prime caused greater Stroop interference on stereotype word trials. By contrast, participants who had formed an implementation intention to ignore the gender of the prime did not show greater Stroop interference on gender-stereotypic words. The same pattern was observed in a second study that included stereotypes of the homeless. Thus, distraction-inhibiting implementation intentions may enhance performance by engaging control over the perception of a stimulus that creates unwanted bias. In this way, they can enhance controlled processing while also reducing the automatic processing of stereotypes.

Response-facilitating implementation intentions refer to a second type of strategy, designed to enhance goal attainment by promoting the execution of a goal-directed action despite any distracting or countervailing influences. For example, Gollwitzer and Brandstätter (1997) conducted an experiment in which university students were asked to form the goal intention of writing a report during the span of a 2-day holiday break (i.e., the 2 days following Christmas Eve). The report was to detail how they had spent their Christmas Eve and was to be sent back to the researchers as soon as possible. Compared with participants in the control condition, those who formed implementation intentions that specified when (e.g., after church on the first holiday) and where (e.g., in a quiet room) to implement their writing task were more likely to return the completed reports by the deadline. Thus, implementation intentions were effective in promoting a planned behavior in response to a situational cue, despite the many holiday distractions. Response-facilitating implementation intentions have also been shown to aid goal attainment in the domain of health behaviors, where goals to engage in preventative behaviors can be derailed by a host of emotional and situational factors. Women who formed a response-facilitating implementation intention to perform breast self-examinations were more likely to do so in the course of a month than those without the strategy (Orbell, Hodgkins, & Sheeran, 1997). Thus, response-facilitating implementation intentions have been shown to improve goal attainment through the reflexive initiation of an intended action, even in the context of powerful biasing influences. In this way, the response-facilitating strategy enhances response control without necessarily affecting automatic processing.

Overview of Present Research

We conducted two experiments to test the effectiveness of distraction-inhibiting (Study 1) and response-facilitating (Study 2) implementation intentions on improving performance on a reaction time measure of implicit race bias. We predicted that implementation intentions would engage reflexive action control, such that task performance would be enhanced despite the lack of deliberation or self-reflection afforded by such tasks. To test our hypotheses, it was critical to use a task that was amenable to the implementation intention procedure in the context of racial stereotypes and that would yield data for computing independent estimates of automatic and controlled components of performance. We chose to use the Shooter Task (Correll, Park, Judd, & Wittenbrink, 2002) because it best met these criteria.

In the Shooter Task, images of Black or White male “targets” appear on the computer screen holding either a gun or a nongun object (e.g., wallet, phone). The participant’s task is to quickly shoot armed targets and to not shoot unarmed targets, via button press, within a short response deadline that precludes deliberation. Implicit stereotyping on this task is characterized by impairments in performance attributable to racial associations, such as the biased tendency to erroneously “shoot” unarmed Blacks more often than unarmed Whites (Correll et al., 2002, 2007). Control in this case is characterized by more accurate performance across trials, despite the potential biasing effects of the target’s race. The design of the Shooter Task is well suited to the use of implementation intentions because cues that are both relevant (i.e., object type) and irrelevant (i.e., target race) to successful goal attainment are presented simultaneously (Gollwitzer & Sheeran, 2006).

To examine the specific processes that are affected by implementation intentions, we used Jacoby’s (1991) process dissociation (PD) procedure for estimating the independent automatic versus controlled patterns of task behavior (as in Payne, 2001). Again, the Shooter Task was appropriate for this analysis because trials are arranged in such a manner that automatic and controlled processes work in concert on some trials (e.g., armed Black trials) and in opposition on other trials (e.g., unarmed Black trials). In the PD framework, the estimate of control (PD-C) represents the probability that one will respond in an accurate, goal-consistent manner, without bias from automatic stereotypes. Both distraction-inhibiting and response-facilitating implementation intentions were expected to increase the role of PD-C in Shooter Task performance. The estimate of automaticity (PD-A) represents the probability that one’s responses will be biased by the racial stereotypes associated with the targets. Because distraction-inhibiting implementation intentions engage a form of control that modifies the perception of a stimulus, they should be associated with a reduced influence of automatic stereotyping processes, as indicated by a decrease in PD-A.

Study 1

The purpose of Study 1 was to examine the effects of distraction-inhibiting implementation intentions. Our main
prediction was that participants using this type of implementation intention strategy on the Shooter Task would perform more accurately across trials, thus showing a lesser degree of implicit stereotyping. We expected that the effect of implementation intentions on performance accuracy would not be associated with slower responding. Finally, because distraction-inhibiting implementation intentions are designed to enhance performance by increasing the control of incoming perceptual information, we predicted that subjects using this strategy would show an increase in PD-C as well as a reduction in PD-A, compared with subjects who did not use an if-then strategy.

**Method**

**Participants.** Seventy-four non-Black, native-English-speaking undergraduates (62% female, 66% White) participated in exchange for course credit. Participants were randomly assigned to the no-strategy or the implementation intention strategy condition.

**Materials.** The Shooter Task was adapted from Correll et al. (2002) and consisted of two blocks of 80 critical trials each, separated by a 30-s rest period, presented on a personal computer using DirectRT and MediaLab (Empirisoft, New York). Each trial began with the presentation of a fixation point, followed by a series of one to four background images (e.g., of hotels, parks) that varied in length from 500 to 800 ms. On the final background image of each series, a Black or White male target holding either a gun or a nongun object appeared (i.e., was superimposed on the background). Participants were instructed to shoot targets carrying guns by pressing a key labeled Shoot and to not shoot targets carrying nongun objects by pressing a different key, labeled Don’t Shoot. Thus, the task design was 2 (race: Black vs. White) × 2 (object: gun vs. no gun). Following previous studies, a response deadline of 630 ms following target onset was imposed to preclude deliberative responding and to elicit a sufficient number of errors to permit reliable analyses of accuracy and PD estimates (Correll et al., 2002; Payne, 2001). A “Too Slow” warning message followed responses that exceeded this deadline.

**Procedure.** Participants were run in groups of up to five in individual computer cubicles. Instructions presented on the computer screen introduced the study as examining perceptual vigilance. All participants read that the objective of the task was to shoot individuals carrying guns by pressing the Shoot key and not shoot those carrying objects by pressing the Don’t Shoot key. Participants in the implementation intention strategy condition received additional instructions stating,

You should be careful not to let other features of the targets affect the way you respond. In order to help you achieve this, research has shown it to be helpful for you to adopt the following strategy: If I see a person, then I will ignore his race!

Participants were instructed to mentally repeat the strategy three times using inner speech and then, after the strategy was removed from the screen, to type it into an open-ended response box to ensure that the strategy was processed.

**Data Reduction**

Given our use of the response deadline, our primary dependent variable was performance, indexed by the error rate across trial types. PD estimates of controlled responding (PD-C) and automatic stereotyping (PD-A) were computed using the formulas described in Payne (2001). PD-C is calculated by subtracting the percentage of incorrect responses on stereotype-incongruent trials from the percentage of correct responses on stereotype-congruent trials \[ C = \frac{P(\text{correct} | \text{congruent}) - P(\text{stereotypic error} | \text{incongruent})}{1 - C} \], and represents the degree to which responses are consistent with the task goal. PD-C estimates are equivalent to accuracy rates but rescaled within the PD framework to range from −1 to 1. PD-A was quantified as the percentage of incorrect responses on stereotype-incongruent trials divided by failures to control \[ A = \frac{P(\text{stereotypic error} | \text{incongruent})}{1 - C} \], and represents the degree to which, when an error is made, the error reflects the influence of racial stereotypes. Thus, PD-A for Black trials represented the extent to which subjects erroneously shot unarmed targets (when control failed), whereas PD-A for White trials represented the extent to which subjects failed to shoot an armed target (when control failed). Additionally, response latencies ranging between 300 and 1,200 ms were natural log-transformed and averaged within trial type for each participant for analysis (but reported in raw milliseconds). Because the “Too Slow” warning appeared only after a response was registered, we were able to include responses that exceeded the deadline in our analysis (as in Amodio et al., 2004).

An a priori decision was made to exclude data from participants who scored at or below chance (50% within trial type; 5 participants). In addition, 1 participant’s mean response latency was an extreme outlier (>3 SDs) and thus excluded. Analyses were conducted using data from 68 participants with valid responses.

**Results**

**Task performance.** Participants’ error rates were submitted to a 2 (race: Black vs. White) × 2 (object: gun vs. no gun) × 2 (strategy: no strategy vs. implementation intention strategy) mixed-factors ANOVA. As a preliminary step, we examined the expected Race × Object interaction to establish that the task produced the expected pattern of implicit stereotyping.
Indeed, this interaction was significant, \(F(1, 66) = 14.21, p < .01\). Simple effects tests revealed that participants were more likely to shoot unarmed Black targets (\(M = .17, SD = .12\)) than unarmed Whites (\(M = .14, SD = .10\)), \(F(1, 66) = 6.72, p = .02\), and more likely to not shoot armed White targets (\(M = .16, SD = .08\)) than armed Blacks (\(M = .12, SD = .08\)), \(F(1, 66) = 12.84, p < .01\).

Next, we examined the main effect of strategy on error rates, which provided a direct test of our primary hypothesis that implementation intentions should improve task performance. This effect was significant, \(F(1, 66) = 7.65, p = .01\), indicating that implementation intention participants made fewer errors overall (\(M = .13, SD = .05\)) than the no-strategy group (\(M = .17, SD = .08\)). This effect was qualified by a significant Race \(\times\) Strategy interaction, \(F(1, 66) = 5.55, p = .02\), which revealed that the effect of the implementation intention (vs. no strategy) was stronger for Black-target trials, \(F(1, 66) = 10.79, p < .01\), than for White-target trials, \(F(1, 66) = 3.15, p = .08\) (see Table 1). A significant Object \(\times\) Strategy interaction also emerged, \(F(1, 66) = 8.94, p < .01\).

Simple effects analyses for this interaction revealed that the implementation intention reduced errors on no-gun trials, \(F(1, 66) = 12.64, p < .01\), but not on gun trials, \(F(1, 66) = .53, p = .47\). The three-way interaction, which would reflect a selective change in performance and thus was not predicted, did not reach significance, \(F(1, 66) = .78, p = .38\). Overall, this pattern supported our prediction that the implementation intention would increase overall performance accuracy.\(^1\) However, this analysis was unable to test our predictions for the specific processes through which distraction-inhibiting implementation intention affected performance. For a more precise test, we used the PD procedure.

**Process dissociation effects.** PD analyses were conducted to examine the mechanisms associated with implementation intention effects. Given the use of the distraction-inhibiting implementation intention in this study, participants in the implementation intention condition were expected to show higher PD-C and lower PD-A estimates relative to the no-strategy participants.

**PD-C estimates.** Participants’ PD-C scores were submitted to a 2 (race: Black vs. White) \(\times\) 2 (strategy: no strategy vs. implementation intention strategy) mixed-factors ANOVA. This analysis produced the predicted main effect of strategy, \(F(1, 66) = 7.65, p = .01\), such that PD-C was greater in the implementation intention condition (\(M = .75, SD = .10\)) than in the no-strategy condition (\(M = .66, SD = .16\)). This effect was qualified by a significant Race \(\times\) Strategy Type interaction, \(F(1, 66) = 5.55, p = .02\). Simple effect analyses revealed that although the implementation intention enhanced control for both targets, the strategy effect was stronger for Black targets, \(F(1, 66) = 10.79, p < .01\), than for White targets, \(F(1, 66) = 3.15, p = .08\) (see Table 2). Although not specifically predicted, this effect is consistent with the notion that implementation intentions should be most effective when self-regulation is difficult, such as in the context of strong racial stereotypes that would interfere with responses on Black-object trials (Gollwitzer & Schaal, 1998; Gollwitzer & Sheeran, 2006). The effect of strategy on the PD-C estimate remained significant when mean response latency was covaried in an analysis of covariance (ANCOVA), \(F(1, 65) = 7.52, p = .01\), suggesting that control did not rely on deliberation.

**PD-A estimates.** PD-A scores were submitted to a 2 (race: Black vs. White) \(\times\) 2 (strategy: no strategy vs. implementation intention strategy) mixed-factors ANOVA. This analysis revealed only a significant interaction, \(F(1, 66) = 4.22, p = .04\). Simple effect analyses indicated that the implementation intention strategy yielded lower PD-A estimates for Black trials compared with the no-strategy condition, \(F(1, 66) = 4.40, p = .04\) (see Table 2). By contrast, PD-A estimates for White trials did not differ as a function of strategy, \(F(1, 66) = 0.77, p = .38\). The strategy effect for Black PD-A estimates remained significant when mean response

### Table 1. Mean Error Rates as a Function of Trial and Strategy Type in Study 1

<table>
<thead>
<tr>
<th>Trial</th>
<th>No strategy</th>
<th>Implementation intention strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>No-gun trials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>.17 (\text{SD} = .11)</td>
<td>.12 (\text{SD} = .08)</td>
</tr>
<tr>
<td>Black</td>
<td>.22 (\text{SD} = .13)</td>
<td>.12 (\text{SD} = .07)</td>
</tr>
<tr>
<td>Gun trials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>.16 (\text{SD} = .08)</td>
<td>.16 (\text{SD} = .09)</td>
</tr>
<tr>
<td>Black</td>
<td>.13 (\text{SD} = .09)</td>
<td>.11 (\text{SD} = .08)</td>
</tr>
</tbody>
</table>

Standard deviations are presented in parentheses. Differences indicate significantly different \((p < .05)\) values within rows on the basis of \(t\) tests.

### Table 2. Process Dissociation (PD) Indices of Automatic and Controlled Response Processes as a Function of Strategy Type in Study 1

<table>
<thead>
<tr>
<th>PD index</th>
<th>No strategy</th>
<th>Implementation intention strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD-Auto</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>.52 (t(22))</td>
<td>.57 (t(23))</td>
</tr>
<tr>
<td>Black</td>
<td>.65 (t(19))</td>
<td>.55 (t(23))</td>
</tr>
<tr>
<td>PD-Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>.67 (t(16))</td>
<td>.72 (t(11))</td>
</tr>
<tr>
<td>Black</td>
<td>.65 (t(18))</td>
<td>.77 (t(12))</td>
</tr>
</tbody>
</table>

PD estimates are presented as probability scores. Standard deviations are presented in parentheses. Differences indicate significantly different \((p < .05)\) values within rows on the basis of \(t\) tests.
latency was covaried in an ANCOVA, $F(1, 65) = 4.61, p = .04$. These results supported the prediction that the distraction-inhibiting implementation intention should reduce the automatic effects of Black stereotypes on the task.

Response latencies. Although analyses on Shooter Task performance focused on error rates given the short response deadline (Correll et al., 2002), an analysis of response latencies was conducted to test our hypothesis that our implementation intention effects did not rely on deliberative processes (indicated by response slowing). A 2 (race: Black vs. White) × 2 (object: gun vs. no gun) × 2 (strategy: no strategy vs. implementation intention strategy) mixed-factors ANOVA on response latencies produced a main effect of object, $F(1, 66) = 379.47, p < .01$, and a Race × Object interaction, $F(1, 66) = 16.18, p < .01$. Simple effects revealed that responses to armed Blacks were faster than to armed Whites, $F(1, 66) = 9.41, p < .01$, and slower to unarmed Blacks than to unarmed Whites, $F(1, 66) = 7.17, p = .01$, across conditions. However, no differences in response latency were found between strategy conditions, $F(1, 66) = .42, p = .52$, indicating that the implementation intention condition did not respond in a more deliberative manner.

Discussion

The findings of Study 1 indicated that the distraction-inhibiting implementation intention strategy increased performance accuracy without the response slowing associated with deliberative forms of control. This finding was consistent with the theory that implementation intentions engage inhibitory control in a reflexive manner. More importantly, Study 1 provided an important test of the processes that underlie the effects of distraction-inhibiting implementation intentions. Analyses of PD estimates supported the theory that distraction-inhibiting implementation intentions improve task performance by increasing inhibitory control and reducing the activation of automatic stereotyping processes. Hence, Study 1 provided critical support for the general theory of implementation intentions and demonstrated their utility for reducing the expression of implicit stereotypes in behavior.

The results of Study 1 complement the findings of Stewart and Payne (2008), who examined whether counterstereotypical implementation intentions could reduce automatic stereotyping effects. Across two studies using the Weapons Identification Task, they found that participants who formed an implementation intention to think “safe” (e.g., versus “quick”) when seeing a Black face did not exhibit the typical Race Prime × Object interaction indicative of race bias on their task. In their studies, counterstereotypical implementation intentions reduced PD-A estimates on trials involving Black faces compared with those involving White faces, whereas control condition participants showed the typical patterns of greater PD-A for Black than for White trials. Although our implementation intention strategy did not focus on a particular racial category, we observed a similar pattern on the PD-A estimates. The “ignore race” strategy was found to reduce automatic processing on responses to Black-target, but not White-target, trials. Given that White is considered to be the normative racial group in American culture, it is possible that participants assumed that the “ignore race” instruction pertained more strongly to Black targets.

An important difference between our study and that of Stewart and Payne (2008) is that we used an implementation intention designed to elicit inhibitory control rather than to alter participants’ mental representations. As a result, we found that our distraction-inhibiting implementation intention strategy also increased PD-C estimates. This difference may be attributed to the fact that our study focused on increasing accuracy (or control), whereas Stewart and Payne’s aimed to reduce response bias caused by automatic associations. Our Study 1 results complement their findings by demonstrating a different process through which implementation intentions may be used to limit the behavioral expression of implicit stereotypes.

Study 2

Having found support for the effects of the distraction-inhibiting type of implementation intention, we turned our attention in Study 2 to examining the response-facilitating type of implementation intention. Response-facilitating implementation intentions engage a form of controlled processing that targets goal-directed action. In the case of the Shooter Task, goal-directed action refers to correct shoot/don’t shoot decisions, despite the distracting influence of race. Hence, as in Study 1, response-facilitating implementation intentions were expected to lead to more accurate performance on the task, which in turn would limit the opportunity for implicit stereotypes to be expressed in behavior. Given the theorized process through which response-facilitating implementation intentions operate, this strategy was expected to enhance controlled processing (PD-C) without affecting automatic processing (PD-A). Again, the effects of implementation intentions were expected to operate reflexively, such that they would not be associated with response slowing.

In addition to testing a different type of implementation intention, Study 2 included a second type of control condition, in which participants used a simple-goal strategy. The simple-goal strategy condition was included to rule out the possibility that any effects associated with the implementation intention strategy could be attributable to the fact that participants were given additional instructions on the task rather than to the specific if-then structure. Because a simple goal does not include the causal link between the situational cue and the intended response, it was not expected to improve performance relative to the no-strategy condition. Participants in the simple-goal condition received the same instructions as those in the no-strategy condition, but the
simple-goal strategy was presented to these participants in a manner that was semantically parallel to how implementation intention participants received their strategy. Hence, our theory-derived a priori predictions regarding the effect of strategy on task performance and control were tested using a set of planned contrasts. The first planned contrast tested the main comparison between the implementation intention condition with the combination of the two control groups, and the second planned contrast compared the two control conditions, which were not expected to differ.

Method

Participants and design. Ninety-two non-Black, native-English-speaking undergraduates (65% female, 56% White) participated in exchange for course credit. Participants were randomly assigned to the no-strategy, the simple-goal strategy, or the implementation intention strategy condition.

Procedure. The procedure was the same as in Study 1, with the exception of the strategies provided to participants in the simple-goal and implementation intention strategy conditions. After reading the objective of the task, these participants once again read that they should not let other factors influence how they respond. To do so, the simple-goal participants were instructed to adopt the following strategy: “I will always shoot a person I see with a gun!” and “If I see a person with a gun, then I will shoot!” Participants in the implementation intention condition read, “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!” As in previous research, the wording of the simple-goal instructions contained the same critical information as that of the implementation intention instructions, with the primary difference being the if-then structure of the strategy.

Data Exclusion and Reduction

Error rates, response latencies, and PD-procedure estimates were computed as in Study 1. We once again excluded any participants who scored below chance; 3 participants were excluded for this reason. Additionally, data from 1 participant were excluded because of extreme outlying response latency scores. Thus, analyses were conducted using data from 88 participants.

Results

Task performance. Error rates were submitted to a 2 (race: Black vs. White) × 2 (object: gun vs. no gun) × 3 (strategy: no strategy vs. simple-goal strategy vs. implementation intention strategy) mixed-factors ANOVA. Replicating Study 1, the Race × Object interaction was significant, F(1, 85) = 11.50, p < .01. Simple effects revealed that participants were more likely to shoot unarmed Black targets (M = .14, SD = .11) than unarmed Whites (M = .12, SD = .09), F(1, 85) = 4.83, p = .03, and more likely to not shoot armed White targets (M = .15, SD = .09) than armed Blacks (M = .13, SD = .08), F(1, 85) = 10.24, p < .01. In addition to the expected task effects, the main effect of strategy from the omnibus ANOVA was marginally significant, F(2, 85) = 2.59, p = .08. However, our a priori predictions were examined directly using planned contrasts, tested within the framework of the omnibus ANOVA. As predicted, the contrast comparing the implementation intention group with the average of the no-strategy and simple-goal strategy groups was significant, t(85) = 2.06, p = .04. This supported our main hypothesis that implementation intention participants (M = .11, SD = .06) would be more accurate across trials than participants not using the if-then strategy (M = .15, SD = .08). The complementary contrast, which compared performance between the no-strategy group and the simple-goal strategy group, was not significant, t(85) = 0.88, p = .38 (see Table 3). Overall, these results supported the hypothesis that a response-facilitating implementation intention (but not a comparable simple goal) would increase performance accuracy, thus limiting the opportunity for implicit stereotypes to be expressed in behavior.

Process dissociation effects. As in Study 1, PD analyses were conducted to examine the mechanisms associated with implementation intention effects (see Table 4). The response-facilitation implementation intention was expected to enhance action control without affecting the perception of race and activation of stereotypes. Therefore, implementation intention participants were expected to show higher PD-C than control participants, but these groups were not expected to differ in PD-A levels.

PD-C estimates. We conducted a 2 (race: Black vs. White) × 3 (strategy: no strategy vs. simple-goal strategy vs. implementation intention strategy) ANOVA on PD-C

### Table 3. Mean Error Rates as a Function of Trial and Strategy Type in Study 2

<table>
<thead>
<tr>
<th>Trial</th>
<th>No strategy</th>
<th>Simple-goal strategy</th>
<th>Implementation intention strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>No-gun trials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>.13</td>
<td>.11</td>
<td>.11</td>
</tr>
<tr>
<td></td>
<td>(.11)</td>
<td>(.08)</td>
<td>(.07)</td>
</tr>
<tr>
<td>Black</td>
<td>.16</td>
<td>.14</td>
<td>.11</td>
</tr>
<tr>
<td></td>
<td>(.14)</td>
<td>(.08)</td>
<td>(.07)</td>
</tr>
<tr>
<td>Gun trials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>.17</td>
<td>.16</td>
<td>.13</td>
</tr>
<tr>
<td></td>
<td>(.09)</td>
<td>(.08)</td>
<td>(.08)</td>
</tr>
<tr>
<td>Black</td>
<td>.15</td>
<td>.13</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>(.09)</td>
<td>(.08)</td>
<td>(.06)</td>
</tr>
</tbody>
</table>

Standard deviations are presented in parentheses. Differing subscripts indicate significantly different (p < .05) values within rows on the basis of pairwise t tests. Note, however, that our critical analysis tested planned contrasts comparing the implementation intention condition with the two control conditions, as described in the text.
Table 4. Process Dissociation (PD) Indices of Automatic and Controlled Response Processes as a Function of Strategy Type in Study 2

<table>
<thead>
<tr>
<th>PD index</th>
<th>No strategy</th>
<th>Simple-goal strategy</th>
<th>Implementation intention strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD-Automatic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>.61</td>
<td>.59</td>
<td>.54</td>
</tr>
<tr>
<td></td>
<td>(22)</td>
<td>(21)</td>
<td>(23)</td>
</tr>
<tr>
<td>Black</td>
<td>.48</td>
<td>.52</td>
<td>.56</td>
</tr>
<tr>
<td></td>
<td>(21)</td>
<td>(18)</td>
<td>(26)</td>
</tr>
<tr>
<td>PD-Control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>.70</td>
<td>.73</td>
<td>.76</td>
</tr>
<tr>
<td></td>
<td>(18)</td>
<td>(12)</td>
<td>(13)</td>
</tr>
<tr>
<td>Black</td>
<td>.69</td>
<td>.73</td>
<td>.80</td>
</tr>
<tr>
<td></td>
<td>(21)</td>
<td>(13)</td>
<td>(11)</td>
</tr>
</tbody>
</table>

PD estimates are presented as probability scores. Standard deviations are presented in parentheses. Differing subscripts indicate significantly different (p < .05) values within rows on the basis of pairwise t tests. Note, however, that our critical analysis tested planned contrasts comparing the implementation intention condition with the two control conditions, as described in the text.

PD estimates. A marginally significant effect of strategy emerged, F(2, 85) = 2.59, p = .08. The contrast that compared the implementation intention condition with the two control conditions was significant, t(85) = 2.06, p = .04, indicating that control estimates were higher for implementation intention participants (M = .78, SD = .12) than for those in the control conditions (M = .71, SD = .15). In addition, an ANCOVA testing the effect of this contrast on the PD-C estimate, in which mean response latency was covaried, remained significant, F(1, 84) = 3.82, p = .05. This finding supported the hypothesis that the implementation intention would enhance reflexive control in task performance, thereby precluding the potential influence of racial stereotypes. The complementary contrast between the two control conditions was not significant, t(85) = 0.88, p = .38, as expected.

PD-A estimates. A 2 (race: Black vs. White) x 3 (strategy: no strategy vs. simple-goal strategy vs. implementation intention strategy) mixed-factors ANOVA on PD-A estimates produced no significant effects for strategy, F < 1; race, F(1, 85) = 2.53, p = .12; or the interaction, F(2, 85) = 1.34, p = .27. The lack of effects for strategy was consistent with the hypothesis that a response-facilitating implementation intention should not affect automatic processing.

Response latencies. To determine whether implementation intention effects occurred reflexively rather than deliberatively, as hypothesized, we examined response latency differences between groups. A 2 (race: Black vs. White) x 2 (object: gun vs. no gun) x 3 (strategy: no strategy vs. simple-goal strategy vs. implementation intention strategy) mixed-factors ANOVA on response latencies produced a main effect of object, F(1, 85) = 141.66, p < .01, such that participants responded more quickly to gun trials (M = 547.25, SD = 53.83) than to no-gun trials (M = 605.47, SD = 49.38) across conditions. A main effect of race also emerged, F(1, 85) = 4.14, p = .05, indicating that participants took longer to respond to White-target trials (M = 578.75, SD = 49.12) than to Black-target trials (M = 573.97, SD = 49.90). Critically, however, response latencies did not vary as a function of strategy, F(2, 85) = 1.14, p = .32, indicating that implementation intentions did not evoke a more deliberative pattern of responding.

Discussion

In Study 2, we tested whether a response-facilitating implementation intention strategy was effective in limiting the expression of implicit stereotypes in behavior. The response-facilitating type of implementation intention was designed to trigger action control, which should enhance performance on the Shooter Task. By increasing response accuracy, there would be less opportunity for stereotypes to be expressed in behavior. Indeed, participants who adopted a task-facilitating implementation intention strategy performed more accurately than those in the control conditions. Furthermore, analyses of PD estimates revealed that this type of implementation intention strategy enhanced controlled processing but did not affect automatic stereotyping processing. Consistent with the theory that implementation intention effects operate reflexively (i.e., without deliberation), the effects of the implementation intention strategy on performance and PD-C estimates were not associated with slowed response latencies.

Whereas the distraction-inhibiting implementation intention used in Study 1 appeared to reduce the automatic processing of racial stereotypes, the response-facilitating implementation intention used in Study 2 affected only controlled processing. Our findings extend those of Study 1 by demonstrating that implementation intentions can still help individuals attain the egalitarian goal of responding without stereotypes, even if the strategy does not directly target the biased associations that may underlie automatic stereotyping effects. Participants in Study 2 were able to respond in a goal-directed manner without any explicit instructions to inhibit or override the influence of racial associations.

The observed results are striking, given that the basic instructions for completing the task were essentially the same for each condition. Indeed, some might be surprised by the fact that the simple-goal strategy did not improve task performance in any way relative to the no-strategy condition, especially because goal intentions have been considered by some theorists to be the most powerful predictor of goal attainment (Ajzen, 1991; Bandura, 1991; Carver & Scheier, 1998; Locke & Latham, 1990). However, forming a goal intention is only the first step in the goal pursuit process and usually does not protect individuals from conflicting attitudinal or behavioral response tendencies (Gollwitzer &
Schaal, 1998). As such, the implementation intention effects observed here appear to be attributable to the if-then structure of the strategy, which allowed participants to directly link an action to the goal-triggering cues presented in the task (see also Oettingen, Hoenig, & Gollwitzer, 2000).

**General Discussion**

The present research examined the effect of implementation intention strategies on the behavioral expression of implicit stereotypes. Across two studies, we found that two different types of implementation intentions were effective in enhancing performance on an implicit race bias task, which in turn limited the expression of implicit stereotypes. Study 1 demonstrated that a distraction-inhibiting implementation intention, which instructed participants to ignore the targets’ race, was associated with an increase in controlled processing and a decrease in automatic stereotyping processes. Study 2 showed that a response-facilitating implementation intention, which focused participants on task-relevant shoot/don’t shoot actions in response to critical stimuli (i.e., guns vs. benign objects), led to an increase in controlled processing without affecting automatic processing. Together, these studies demonstrate the effectiveness of two different types of implementation intentions for reducing the influence of race on behavior without the need for deliberation. In what follows, we discuss the implications of these findings for theories of control in social psychology and for approaches to reducing the behavioral expression of intergroup bias.

**Reducing Expressions of Intergroup Bias Through Reflexive Control**

The present work represents a new approach to reducing implicit stereotyping effects that focuses on the reflexive engagement of control rather than on attempting to directly alter latent forms of bias built up during the course of a lifetime. Although changing biased associations is a praiseworthy goal, the current findings suggest that the quick interventions typically used in psychological experiments may be more effective in modulating behavioral responses or the temporary accessibility of stereotypes than in undoing highly edified knowledge structures. Given the pervasiveness of stereotypes in the social environment, a focus on enhancing reflexive control is a practical short-term strategy that may result in the long-term change of social perceptions and experiences that are necessary to eventually break down biased associations (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002).

The effects of the two types of implementation intentions examined here suggest two different ways that a prejudice reduction intervention may affect the expression of stereotypes. The results from Study 1 suggest that the engagement of inhibitory control, by means of a distraction-inhibiting implementation intention, may alter initial perceptions of task stimuli in a way that preempts the activation of race and its associated stereotypes. That is, the targets of control are perceptual processes, which should thus limit the activation of racial stereotypes and their expression in behavior. By comparison, the results from Study 2 suggest that the engagement of action control, by means of a response-facilitating implementation intention, promotes goal-directed task behavior despite the activation of racial stereotypes, thereby limiting their implicit influence on task performance. Here, the target of control is behavior. An important contribution of the present work is that it specifies how a reflexive form of control can target two different processes related to the behavioral expression of implicit stereotypes: initial perceptions of racial stimuli and goal-driven behavior.

It is notable that in psychometric terms, these strategies are designed to enhance accuracy without necessarily affecting bias. That is, a person may still have a tendency to associate Black people with violence and thus be more likely to shoot unarmed Blacks than to shoot unarmed Whites. However, the increase in response accuracy afforded by the implementation intentions used here would reduce the expression of such biases in behavior. If accuracy is extremely high, than an underlying bias would not be expressed, and the net result is a reduction in racial discrimination.

**New Questions and Future Directions**

Like any new finding, our results raise several new questions. For example, can implementation intentions have enduring effects on the expression of stereotypes over time? Previous investigations of implicit prejudice interventions have shown effects lasting a day or two (e.g., Dasgupta & Greenwald, 2001; Olson & Fazio, 2006), but research examining the longevity of implementation intention effects on health behavior suggests that they may last weeks (e.g., Gollwitzer & Brandstätter, 1997) and even months (e.g., Orbell et al., 1997). Why might an implementation intention strategy have longer-lasting effects than previous methods of reducing prejudiced behavior? As previously discussed, implementation intentions serve to establish a plan, which may then be carried out reflexively in future situations in which racial cues are present (see also Monteith, 1993; Monteith, Ashburn-Nardo, Voils, & Czopp, 2002; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999). By comparison, many previous interventions have attempted to change underlying representations of racial groups, which may be more difficult to maintain upon reexposure to societal stereotypes outside the laboratory.

Another question concerns whether implementation intention effects generalize to different situations and types of behaviors. The present studies (as well as those by Stewart & Payne, 2008) examined the effects of implementation intentions on a circumscribed measure of implicit race bias within a tightly controlled experimental context. As such, it
is unclear whether the specific implementation intention strategies used here would generalize to other situations, such as real-life interactions. However, the phrasing of an implementation intention can be modified to pursue the broader goal of egalitarianism in any particular situation through a specific response. For example, an implementation intention may be phrased to initiate more controlled (i.e., careful) responding in a general sense whenever a racial cue is encountered (Monteith, 1993; Monteith et al., 2002). Future research is needed to explore the effectiveness of implementation intentions tailored to specific intergroup situations and responses as well as those that engage a more general form of controlled processing when a racial cue is encountered.

Implications for Implementation Intention Theory

Although the main purpose of the present work was to test the ability of implementation intentions to limit the behavioral expression of implicit stereotypes, the results also have important implications for implementation intention theory. A key component of implementation intention theory is that if-then plans operate through reflexive, rather than reflective, processes. Although this idea is supported by findings of several other studies (e.g., Bayer, Achtzheimer, Gollwitzer, & Moskowitz, 2009; Brandstätter et al., 2001; Schweiger Gallo, Keil, McCulloch, Rockstrob, & Gollwitzer, 2009), the present studies provide unique support for this position by showing that the implementation intentions tested here led to increases in PD-C without a slowdown in responding that typically characterizes traditional deliberative mechanisms of control. That is, although deliberation may be necessary for the formation of an if-then plan, it is not required for the plan’s implementation. The present work also demonstrated that distraction-inhibiting implementation intentions affect both controlled and automatic processing, whereas response-facilitating implementation intentions target only controlled processing (which promotes goal-directed behavior by overriding automatic influences). Hence, the present studies provide an important advance in research on implementation intentions.

Conclusion

It is commonly assumed that expressions of implicit stereotyping may be reduced only through effortful deliberation or changes in one’s cognitive representations of a racial outgroup. To the contrary, the present research suggests that the influence of implicit stereotypes on behavior can be limited through the use of strategies that engage reflexive forms of control. These findings elucidate the process through which implementation intentions promote goal-directed behavior and add to the growing repertoire of strategies for reducing prejudice.

Acknowledgment

The authors thank Joshua Correll for providing the Shooter Task stimuli and instructions for use in these studies.

Declaration of Conflicting Interests

The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The authors received the following financial support for the research and/or authorship of this article: research supported in part by National Science Foundation grant BCS 0847350 to David M. Amodio.

Note

1. We define control in the traditional manner as the successful implementation of a goal-driven response, which may be operationalized as performance accuracy. According to this general model of control, an increase in control should be indicated here by a main-effect increase in performance accuracy across trials on a task (e.g., Jacoby, 1991). By comparison, some research on prejudice and stereotyping has characterized “prejudice control” as a selective reduction or reversal in the stereotyping or derogation of an outgroup target without any effect on other responses (i.e., a three-way interaction in the Shooter Task). However, this selective form of control is not consistent with domain-general models of control in psychology (reviewed by Amodio & Devine, 2010) that emphasize that control corresponds to one’s focal task goal (e.g., to make correct shoot/don’t shoot judgments) rather than an implied and/or tacit secondary goal of selectively suppressing racial bias in reaction time tasks.

References


