tapraid5/z2g-perpsy/z2g-perpsy/z2g00215/z2g4288d15z | xppws | S=1 | 1/7/15 | 3:54 | Art: A-2013-0804

Toward a Comprehensive Understanding of Executive Cognitive Function in Implicit Racial Bias

Tiffany A. Ito and Naomi P. Friedman
University of Colorado Boulder

Bruce D. Bartholow
University of Missouri

Joshua Correll, Chris Loersch, Lee J. Altamirano, and Akira Miyake
University of Colorado Boulder

Although performance on laboratory-based implicit bias tasks often is interpreted strictly in terms of the strength of automatic associations, recent evidence suggests that such tasks are influenced by higher-order cognitive control processes, so-called executive functions (EFs). However, extant work in this area has been limited by failure to account for the unity and diversity of EFs, focus on only a single measure of bias and/or EF, and relatively small sample sizes. The current study sought to comprehensively model the relation between individual differences in EFs and the expression of racial bias in 3 commonly used laboratory measures. Participants (N = 485) completed a battery of EF tasks (Session 1) and 3 racial bias tasks (Session 2), along with numerous individual difference questionnaires. The main findings were as follows: (a) measures of implicit bias were only weakly intercorrelated; (b) EF and estimates of automatic processes both predicted implicit bias and also interacted, such that the relation between automatic processes and bias expression was reduced at higher levels of EF; (c) specific facets of EF were differentially associated with overall task performance and controlled processing estimates across different bias tasks; (d) EF did not moderate associations between implicit and explicit measures of bias; and (e) external, but not internal, motivation to control prejudice depended on EF to reduce bias expression. Findings are discussed in terms of the importance of global and specific EF abilities in determining expression of implicit racial bias.

Keywords: implicit bias, inhibition, working memory, task-switching, automatic and controlled processing

Supplemental materials: http://dx.doi.org/10.1037/a0038557.supp

Borrowing from paradigms initially designed to demonstrate semantic associations between word pairs (e.g., Collins & Loftus, 1975; Neely, 1977; Posner & Snyder, 1975), social psychologists discovered 30 years ago that reaction time (RT) and accuracy based laboratory measures could be used to assess cognitive associations between social categories, like race and gender, and attributes associated with those categories, such as stereotypes and evaluations (e.g., Dovidio, Evans, & Tyler, 1986; Gaertner & McLaughlin, 1983). The prospect that this mental content could be assessed without reliance on self-report was met with great enthusiasm, particularly as it was becoming clear at the time that self-reported intergroup attitudes were artificially positive, masking an underlying, stubborn basis of prejudice (Crosby, Bromley, & Saxe, 1980; Devine & Elliot, 1995). In essence, these measures promised a way to assess stereotype-based associations that people were increasingly unwilling (or unable) to report, thereby circumventing well documented limitations inthe willingness and ability to accurately introspect on the cognitive processes that produce overt behaviors (Kunda, 1990; Nisbett & Wilson, 1977; Zajonc, 1980).

Given their promise for assessing associations that exist outside of conscious awareness and control (Devine, 1989; Greenwald & Banaji, 1995), this new class of measures was initially viewed as a pure reflection of the existence (and, in some cases, the strength; see Fazio, Jackson, Dunton, & Williams, 1995) of underlying, "automatically" activated associations. That is, they were welcomed as measures of

Tiffany A. Ito, Department of Psychology and Neuroscience, University of Colorado Boulder; Naomi P. Friedman, Department of Psychology and Neuroscience and Institute for Behavioral Genetics, University of Colorado Boulder; Bruce D. Bartholow, Department of Psychological Sciences, University of Missouri; Joshua Correll, Chris Loersch, Lee J. Altamirano, and Akira Miyake, Department of Psychology and Neuroscience, University of Colorado Boulder.

This work was supported by National Science Foundation Grant BCS 0847872 and NIH Grant MH063207. We are grateful to Geoff Kerr, Lindsey Newnes, and Christopher R. Engelhardt for assistance with data collection.

Correspondence concerning this article should be addressed to Tiffany A. Ito, Department of Psychology and Neuroscience; University of Colorado; 345 UCB; Boulder, CO 80309-0345. E-mail:tiffany.ito@colorado.edu

 0022-3514/15/$12.00 http://dx.doi.org/10.1037/a0038557

processes that occur quickly, without intention, and with minimal demand on cognitive resources, unlike previously developed explicit measures.[1] These new procedures—now referred to as "implicit" measures—were assumed to be free from distortions imposed by more reflective, effortful controlled processes (Bargh, 1999; De Houwer et al., 2009; Fazio et al., 1995; Greenwald & Banaji, 1995; Greenwald, McGhee, & Schwartz, 1998). This characterization has been described as *task dissociation*, assuming that such response-time- or accuracy-based measures solely reflect automatic associations, in contradistinction to explicit self-report measures, which were assumed to be influenced by controlled processes. More recent perspectives, however, have begun to reject this approach in favor of a *process dissociation* view. Process dissociation posits that all overt behaviors—whether pressing a button on a computer or completing a questionnaire—reflect the influence of both automatic and controlled processing components (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Jacoby, 1991; Payne, 2001, 2005). The purpose of the present research is to provide an extensive, methodologically rigorous test of the process-dissociation perspective by clarifying the role of cognitive control in implicit measures of racial bias (from here on, *implicit bias tasks*).

The Role of Control in the Expression of Implicit Racial Bias

With the process dissociation perspective as a starting point, a number of researchers have argued that, because implicit bias tasks require participants to make overt behavioral responses (e.g., button pressing), performance must reflect the influence of controlled processes to some extent (see Amodio & Mendoza, 2010; Conrey et al., 2005; Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007; Payne, 2001, 2005; Sherman, Klauer, & Allen, 2010). Consider the frequently used Weapon Identification Task (WIT; Payne, 2001) in which racial category cues (faces of White and Black men) are presented just prior to briefly presented target objects that participants must classify as guns or tools (by pressing one of two keys) prior to a rapid response deadline. On some trials—for example, when Black faces precede gun targets—this task can be performed accurately by either careful processing of the target object or by relying on automatically activated stereotypical associations linking young Black men with gun violence. On such *stereotype congruent* trials, both the automatically activated stereotype and more controlled explicit task goals facilitate the same response (i.e., pressing the "gun" key). In contrast, other trials, such as those in which Black faces precede tool targets, require participants to overcome the influence of the stereotype to make the correct response. On these *stereotype incongruent* trials, automatic and control-related processes call for opposing responses, and the correct response depends upon exertion of control. A wide variety of racial bias tasks, such as the Implicit Association Test (IAT; Greenwald et al., 1998) and First Person Shooter Task (FPST; Correll et al., 2002), follow a similar conceptual structure.

Mathematical models using Jacoby's (1991) process dissociation procedure (PDP; Payne, 2001), or more complex multinomial modeling procedures (Conrey et al., 2005), which separately quantify the contribution of automatic and controlled components to behavior, consistently demonstrate that controlled processes substantially contribute to performance on implicit bias tasks. In the PDP analysis, estimates for automaticity and control reflect error rates on stereotype-congruent and stereotype-incongruent trials. Specifically, PDP control is computed as the proportion of congruent trials on which participants respond correctly minus the proportion of incongruent trials on which they commit a stereotype-related error; PDP automatic is computed as the proportion of threat responses on nonthreatening trials divided by the quantity (1 − PDP control), which represents the likelihood of an incorrect threat response when control fails. (The full set of PDP equations can be found in Payne, 2005.) Using these formulas, performance on a variety of implicit measures has been shown to depend not just on "automatic" processes, but also on participants' ability to exert control over the influence of stereotypic associations on behavioral responses (e.g., Amodio, Devine, & Harmon-Jones, 2008; Amodio et al., 2004; Bartholow, Henry, Lust, Saults, & Wood, 2012; Conrey et al., 2005; Payne, 2001, 2005; Payne, Lambert, & Jacoby, 2002; Sherman et al., 2008). Such results clearly challenge the idea that performance on these measures simply reflects automatic associations.

Despite their potential implications, such findings do not directly address the extent to which individual differences in control-related abilities in general contribute to performance on implicit bias measures. In other words, the output of such models represents the engagement of task-specific control processes, which should be influenced by individual differences in control-related abilities but do not themselves reflect solely those abilities. The mathematical approaches may also be limited because behavioral responses from a single task are used for multiple purposes: to assess racial bias, to estimate the influence of automatic processes, and to estimate the influence of cognitive control. This leaves open the possibility that control and racially biased behavior are related in these studies simply because they draw on the same responses.

Other research has endeavored to separately assess bias and control-related abilities by administering racial bias tasks in conjunction with separate measures of executive functioning (EF; Amodio et al., 2008; Klauer, Schmitz, Teige-Mocigemba, & Voss, 2010; Payne, 2005; Siegel, Dougherty, & Huber, 2012). Across these studies, participants demonstrating stronger cognitive control abilities showed less bias in task performance. Other investigators have used psychophysiological measures of processes arguably reflective of EF, recorded while participants complete implicit bias tasks, to augment behavioral assessments, and have found that participants whose neural responses indicate greater implementation of cognitive control show less bias in their behavior (e.g., Amodio et al., 2008, 2004; Bartholow, Dickter, & Sestir, 2006; Bartholow et al., 2012; Correll, Urland, & Ito, 2006; see also Beer et al., 2008; Richeson et al., 2003).

Perhaps the most convincing data on the topic comes from studies in which participants' ability to exert control during implicit measures is directly manipulated. For example, Payne, Lam-

Fn1 [1] The term "automatic" is used widely, but is often poorly defined (cf. Bargh, 1994). Automatic processes typically incorporate one or more of the following components: speed, lack of intention, lack of awareness, minimal demand on cognitive resources, parallel processing, unavoidability (i.e., difficulty overriding the process once it is initiated), and consolidation of subordinate elements. In this article we use the term chiefly to refer to processes that are unintentional and difficult to avoid. We suggest, however, that much as this paper draws meaningful distinctions regarding different forms of "control," the field would benefit from a more nuanced treatment of "automaticity." Neither construct is monolithic.

bert, and Jacoby (2002) manipulated how quickly participants were required to respond in the WIT. Time pressure increased the amount of bias expressed on the task, and a PDP analysis showed that the shift occurred not as a function of changes in the influence of automatic processes but because the manipulation reduced participants' ability to exert control over their behavior (see also Conrey et al., 2005). Similar effects have been found with other manipulations of cognitive control, including depleting control by forcing participants to exert cognitive effort immediately before the implicit task (Govorun & Payne, 2006), introducing an additional cognitively demanding task while measuring racial bias (i.e., increasing cognitive load; Schmitz, Teige-Mocigemba, Voss, & Klauer, 2013), increasing the salience of racial prejudice prior to measuring implicit bias (Siegel et al., 2012), and having participants consume alcohol prior to completing the implicit measure (Bartholow et al., 2006, 2012; Schlauch, Lang, Plant, Christensen, & Donohue, 2009).

Unresolved Issues

Although converging evidence based on a range of approaches suggests that performance on implicit measures of racial bias depends on more than automatically activated associations in memory, investigations of these issues to date have proceeded in a somewhat piecemeal fashion, leaving two core issues unresolved. First, and of critical theoretical importance, there is currently no unifying framework for understanding the specific processes through which control is implemented in the context of implicit racial bias. Within the literature, controlling implicit racial bias is often characterized as involving intention, effort, and limited capacity resources (e.g., Bartholow et al., 2006; Conrey et al., 2005; Devine, 1989; Lambert et al., 2003; Payne, 2001, 2005; Payne, Shimizu, & Jacoby, 2005; von Hippel, 2007), but while these features may accurately describe the conditions under which controlled processes operate, they do not define the qualitative nature of the processes themselves (i.e., what a particular process does). Even definitions of control that emphasize goal-congruent processing or pursuing distal over proximal goals (e.g., Baumeister, Vohs, & Tice, 2007; Fujita, 2011) do not identify the concrete processes needed to achieve these outcomes.

Within the stereotyping and prejudice literature, a number of unique mechanisms have been suggested, including inhibiting the impact of automatically activated stereotypic associations and activation of more egalitarian associations and/or response tendencies (Devine, 1989; Macrae, Bodenhausen, Milne, & Jetten, 1994; Mendoza, Gollwitzer, & Amodio, 2010; Monteith, Ashburn-Nardo, Voils, & Czopp, 2002; Monteith, Sherman, & Devine, 1998; von Hippel, Silver, & Lynch, 2000). Here we argue that the processes typically attributed to "control" in the stereotyping and prejudice literature overlap with what are considered in the cognitive science literature to be EFs (see also Klauer et al., 2010; Payne, 2005; Richeson & Shelton, 2003). This view is in line with previous demonstrations that control over race-based responses relies on the same neural conflict monitoring and control adjustment processes identified in cognitive neuroscience models of control more generally (e.g., Amodio et al., 2008, 2004; Bartholow et al., 2006, 2012; Correll et al., 2006). Moreover, the current work goes beyond such previous demonstrations in that, rather than adopting a monolithic and undifferentiated view of EF, we build on the latest advances in cognitive science indicating that EF is a multifaceted construct consisting of processes including response inhibition, working memory updating, and task switching. These aspects of EF have been shown to possess both "unity and diversity" (see Teuber, 1972) in a variety of populations, including young adults, college students, children, and older adults (e.g., Fisk & Sharp, 2004; Friedman et al., 2008; Huizinga, Dolan, & van der Molen, 2006; Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003; Miyake et al., 2000). Unity refers to the fact that these separate EFs significantly correlate with each other; diversity refers to the fact that the correlations are not total, indicating important differences among the different types of EF.

This pattern of unity and diversity has had important implications for understanding the relations of EFs to other constructs, as the apparent influence of EF depends on the constructs or specific measures examined (e.g., Friedman et al., 2008, 2011; Miyake et al., 2000; Willcutt et al., 2001). For example, general intelligence is most highly related to working memory updating ability (Friedman et al., 2006), whereas deficits in behavioral and attentional control are more highly related to response inhibition (Friedman et al., 2007; Young et al., 2009). If we are to understand complex behavioral outcomes like the expression of racial bias, treating control as a unitary construct is insufficient (Klauer et al., 2010). EF is a nuanced and multifaceted construct, and forming a full understanding of how control over automatic racial associations is implemented necessitates evaluating multiple aspects of EF and their relationships with implicit racial bias.

Another issue warranting further investigation is the degree to which relations of EF to implicit bias are similar across different bias tasks, or whether unique task structures employed by different measures will lead them to differentially relate to the various components of EF described above. Addressing this requires not only multiple measures of EF, but also implicit bias. In addition to facilitating consideration of how multiple aspects of EF relate to multiple measures of EF, the use of multiple tasks also provides various measurement advantages. From a task impurity perspective, any given measure is a fallible indicator of the underlying construct of interest, reflecting both construct-specific and task-specific variability (as well as random error; e.g., Blanton, Jaccard, Gonzales, & Christie, 2006; Burgess, 1997; Cunningham, Preacher, & Banaji, 2001; Fazio & Olson, 2003 Miyake et al., 2000; Phillips, 1997). This measurement problem is exacerbated by the poor reliability of many implicit bias and EF tasks (Cunningham et al., 2001; Denckla, 1996; Rabbitt, 1997). Although some studies have examined the role of cognitive control in implicit measures, many have tended to use a single measure of either implicit bias (e.g., Klauer & Mierke, 2005; Mierke & Klauer, 2003; Siegel et al., 2012) or control/EF (e.g., Payne, 2005; Klauer & Mierke, 2005; Mierke & Klauer, 2003; but see Amodio et al., 2008; Klauer et al., 2010).

To further illustrate the importance of multiple measures of both EF and bias, consider the differing results obtained in recent studies. The first example comes from studies by Payne and colleagues examining performance on implicit bias tasks, including the WIT and an evaluative priming task (Govorun & Payne, 2006; Payne, 2005). In each study, one aspect of EF—inhibitory ability—was measured by a single task (either the Stroop or antisaccade task), and the authors found that EF task performance was positively associated with estimates of control from PDP

tapraid5/z2g-perpsy/z2g-perpsy/z2g00215/z2g4288d15z | xppws | S=1 | 1/7/15 | 3:54 | Art: A-2013-0804

APA NLM

analyses. These results seem to suggest that (at least some) bias measures are sensitive to participants' *inhibitory ability*. In contrast, in one of the most methodologically and analytically sophisticated studies to date, Klauer et al. (2010) examined the relationship between performance on the IAT and performance on measures of inhibition, working memory, and switching. This work took a latent variable approach, employing multiple IATs tapping different constructs and multiple measures of each aspect of EF. Klauer et al. (2010) found only a negligible relationship between inhibitory ability and bias, when controlling for variance in the other EFs. Performance on the IAT was instead related to switching ability. Though Klauer et al.'s (2010) work focused purely on construct-neutral IAT method variance (i.e., not exclusive to race bias), these studies suggest that performance on different bias tasks reflects contributions from distinct EF abilities. If EF demands vary across different implicit bias measures, relations observed between a given measure of EF and any particular bias task cannot be assumed to generalize to other measures. Hence, comprehensive understanding of these relationships requires the use of multiple measures of both EF and bias.

Current Study

Based on these considerations, we conducted a large-scale, systematic study of the relation between EF and implicit racial bias that builds on past research but also expands it in several important ways. First, the current work represents a combination of cutting edge approaches in assessment of EF (from cognitive science) and racial bias (from social cognition), along with sophisticated quantitative modeling techniques. These innovations allow us to evaluate the degree to which distinct facets of EF, assessed at the level of latent variables, are involved in the expression of implicit racial bias. In line with the EF research described above, we employed a set of nine EF tasks, allowing independent assessments of response inhibition, working memory updating, and task switching at the latent variable level. Following Miyake and Friedman (2012), we incorporate a nested factors model that more directly captures the unity and diversity of these EFs. In this model, unity is captured by a "Common EF" factor on which all EF tasks load, and diversity is captured by two additional factors ("Updating-specific" and "Shifting-specific") on which the working memory updating and task-switching measures load, respectively. These latter two factors capture the correlations among the updating and switching tasks that are not captured by the Common EF factor (i.e., variance that makes these factors distinct from each other and from Common EF). Notably, although research on the control of implicit racial biases often implicates inhibitory processes (e.g., inhibiting the influence of an undesirable behavioral tendency; see Devine, 1989; Mendoza et al., 2010; Monteith et al., 2002; von Hippel, Silver, & Lynch, 2000), the current model lacks an "Inhibition-specific" factor because inhibition is represented by the Common EF factor. That is, once variance common to all three types of EF measures is accounted for by estimation of the Common EF factor, there are no remaining correlations among the inhibition tasks to account for, a pattern that has been replicated in multiple independent datasets (see Miyake & Friedman, 2012), including that from the current study.

This nested factors model allows our results to be interpreted in terms of the proposed mechanisms that are common and unique to these EFs (Friedman et al., 2011; Miyake & Friedman, 2012). Specifically, Miyake and Friedman (2012) proposed that individual differences in Common EF tap the ability to actively maintain task goals, particularly in the face of interference, and use these goals to direct ongoing processing. This ability is key to all EF tasks and may be particularly important in response inhibition tasks (Munakata et al., 2011). In contrast, they proposed that Shifting-specific ability may be more related to the ability to quickly let go of these goals when necessary, to flexibly adapt ongoing behavior to changing situational demands. Updating-specific abilities may reflect efficient gating of working memory, as well as episodic retrieval. Assessment of these individual aspects of EF allows us to examine the degree to which qualitatively different aspects of EF contribute to the operation of control within the context of implicit racial bias.

A second important advance made in the current study is the inclusion of three widely used implicit measures of racial bias: the WIT, FPST, and IAT. The use of multiple indices of bias in the same participants has two distinct advantages. First, this approach permits examination of the degree of association among the bias measures (e.g., Cunningham et al., 2001). Importantly, we modified the implicit measures to assess the same basic aspect of racial stereotypes of African Americans (the association between African Americans and danger). In doing so, we eliminate assessment of different automatic associations as one possible reason for the dissociations among the tasks reported in prior research (e.g., Brauer, Wasel, & Niedenthal, 2000; Fazio & Olson, 2003; Sherman, Rose, Koch, Presson, & Chassin, 2003). We can then assess whether any remaining discrepancy between implicit bias tasks reflects either the differential contribution of cognitive control (in general) to a given measure of bias or the specific relationship between one particular measure of bias and one particular subcomponent of EF. Here, we focused on the African American–danger association because of its widespread assessment in past research (across multiple implicit measures) and the consistent finding that this stereotype exists broadly across different parts of American society (Chen & Bargh, 1997; Oliver, 2003). Second, using multiple measures improves assessment of the underlying construct of interest by circumventing task impurity problems (i.e., task-specific method variance; see Mierke & Klauer, 2003, for an example involving the IAT). We thus take a latent variable approach to modeling both EF and implicit bias, allowing us to extract overarching factors representing the variables of interest from our large battery of tasks and measures.

The third goal of the current study was to provide a definitive data set that was largely immune to the criticisms levied at past research in this area. To accomplish this, we incorporated a number of basic design improvements into our study. Perhaps the most important is the relatively large number of participants included in this study ($N = 485$), which provides the power necessary to test nuanced predictions concerning the strength of correlations among our measures. Also, participants were recruited from three universities located in distinct geographical and cultural areas of the U.S. Together, these factors make this one of the largest and most heterogeneous laboratory-based studies of racial bias to date, especially on the role of cognitive control in implicit bias. Furthermore, each participant completed a two-session protocol involving assessment of EFs during Session 1, and assessment of implicit bias during Session 2. These sessions were separated by roughly a week to ensure that any relationships between the two classes of measures were due to persistent individual differences that were stable over time.

In addition to comprehensively assessing EFs and implicit bias, we administered several questionnaire measures designed to assess racial attitudes in a more explicit fashion. Inclusion of these explicit measures allowed us to test several additional questions concerning the relationship between implicit and explicit bias measures and relations between EF and individual differences in motivation to control prejudice. These measures were also assessed during Session 1 so that any relationships between the implicit and explicit measures of racial bias would be due primarily to stable individual differences. Critically, our large sample size allows us to address these questions within a well-powered design, thereby eliminating a substantial limitation found in most other studies examining EF and bias. Given the scope and complexity of the research questions investigated here, and to provide an organizing conceptual framework for the study the next section briefly outlines our primary (H1 & H2) and ancillary (H3–H6) hypotheses.

Primary Hypotheses

Our primary hypotheses and the fundamental motivation for this research concern the degree to which EF and automatically activated stereotypic associations affect participants' performance on standard, laboratory-based measures of implicit bias. H1 and H2 both focus on implicit bias as the dependent variable of interest.

Hypothesis 1 (H1): Implicit bias reflects both EF and automatic processes. Our primary question involves the degree to which performance on implicit measures of racial bias reflects variability in both EF abilities and automatic processes linking race and danger. Using EF performance as a reflection of the degree to which executive control processes can be brought to bear when completing an implicit bias task, and the PDP estimates of automatic bias as a measure of the contribution of automatic processes, we expect both EF and PDP automatic estimates (from here on *PDP auto*) to directly predict racial bias in performance (from here on *performance bias*). Specifically, higher scores on EF tasks should be associated with reduced performance bias on the WIT, FPST, and IAT, whereas higher PDP auto should be associated with greater performance bias on these tasks.

Hypothesis 2 (H2): EF moderates the effect of automatic processes on implicit bias. In addition to predicting direct effects of both EF and automatic processes on performance bias, we also expect them to interact. To the degree that bias expression depends on the ability to exert control over the influence of automatically activated associations on behavior, EF should moderate the relation between PDP auto and performance bias (Payne, 2005). Specifically, automatic processes should more directly drive performance bias among those with weaker EF abilities. By the same logic, automatic processes should be a weaker predictor of performance bias among those with stronger EF abilities.

Critically, both hypothesis H1 and H2 will be assessed using the unity and diversity nested factors EF model (Miyake & Friedman, 2012). This model will allow us to further test whether particular aspects of EF are differentially implicated in these predicted relations.

Ancillary Hypotheses

In addition to these primary hypotheses, the relatively large scale of the current investigation provides the opportunity to test several ancillary hypotheses concerning relationships among implicit measures, explicit measures, EF, and motivational factors.

Hypothesis 3 (H3): EF is related to PDP control but not auto. Within the process dissociation framework, computations of controlled processes are thought to reflect the operation of EFs whereas estimates of automatic processes are not (Payne, 2005; see also Govorun & Payne, 2006). Although the differential relationship of EFs with PDP control and auto has been supported empirically (Govorun & Payne, 2006; Payne, 2005, see also Amodio et al., 2008, 2004; Lambert et al., 2003; Mendoza et al., 2010; Stewart & Payne, 2008), prior studies have been limited both in the types and number of EF and/or implicit bias measures assessed. The present study allows us to examine this relationship in a more comprehensive manner through use of the unity and diversity EF model. Based on past studies finding an association between PDP control and measures related to Common EF (Govorun & Payne, 2006; Payne, 2005), we expect a relationship between PDP control and Common EF. To the degree that controlling the influence of automatic stereotypic associations in these tasks requires flexible switching between different goals and behaviors (which draws on Shifting-specific ability), and/or the active manipulation of information in working memory (which draws on updating-specific ability), PDP control may also relate to Shifting-specific and Updating-specific abilities.

Hypothesis 4 (H4): EF moderates the relation between implicit and explicit bias. Working from a classic dual process perspective, weak correlations between implicit and explicit measures could be assumed to reflect the greater influence of controlled processes on explicit than on implicit measures (e.g., Blair, 2001; Cunningham et al., 2001; Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002; Fazio et al., 1995; Fazio & Olson, 2003; Greenwald, McGhee, & Schwartz, 1998; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Hofmann, Gschwendner, & Schmitt, 2005). That is, while both implicit and explicit responses could draw on an individual's beliefs and feelings about a group, explicit responses might be more influenced by additional considerations such as motivation to control bias. Consistent with this assumption, a higher correspondence between the two types of measures has been obtained when situational pressures to implement control are minimized (Payne, Burkley, & Stokes, 2008; Ranganath, Smith, & Nosek, 2008 see also Hofmann et al., 2005; Koole, Dijksterhuis, & van Knippenberg, 2001). Here we test an individual difference analog of this effect. Given our assumption that control depends on EF, we predict that EF moderates the relationship between implicit and explicit measures of racial bias (see also, Payne, 2005), with the strength of the implicit–explicit relationship increasing for those who have less capacity to modify their explicit responses (i.e., as EF decreases; Payne, 2005).

Hypothesis 5 (H5): Bias is influenced by motivation to control prejudice. Motivation to control prejudiced responses has

APA NLM
tapraid5/z2g-perpsy/z2g-perpsy/z2g00215/z2g4288d15z | xppws | S=1 | 1/7/15 | 3:54 | Art: A-2013-0804

been argued to moderate the degree to which stereotypical associations impact behavior (Devine, 1989; Dunton & Fazio, 1997; Monteith, 1993; Plant & Devine, 1998). We thus predict that participants who are motivated to control the expression of prejudice will show less implicit and explicit bias, even when controlling for the influence of EF. This may be especially true of motivation that derives from internal, personally held egalitarian beliefs. People for whom control over bias derives from a desire to conform to perceived external normative standards may actually show higher levels of bias relative to those with lower levels of external motivation (Amodio et al., 2008, 2003; Devine et al., 2002; Gonsalkorale et al., 2011).

Hypothesis 6 (H6): EF moderates the effect of motivation to control prejudice on bias. Finally, in an extension of H5, we examine the degree to which motivation to control prejudice interacts with EF to predict implicit and explicit bias. While domain-general EF ability may be necessary to translate regulatory intentions into behavior (as assessed under H1 and H2), the specific concern of expressing racial bias should also be critical for recruiting such control (cf., Amodio et al., 2008). Thus, although concern about not appearing racially biased could motivate individuals to control their behavior, successfully implementing control may depend on cognitive capacity. If so, motivation to control prejudiced responses and EF should interact in predicting racial bias such that the impact of high motivation to control prejudice is most pronounced among participants with high EF abilities. This hypothesis is consistent with general theorizing about the interactive effects of motivation and ability (e.g., Fazio, 1990). Although the independent effects of domain-specific motivation and domain-general EF have been assessed before (Payne, 2005), their interactive effects have not to our knowledge been evaluated. Again, given that internal sources of motivation appear to translate more readily into lower bias, high EF may be particularly beneficial with high internal as compared to external motivation.

Method

Participants

Four-hundred and eighty five undergraduate participants (246 men) were recruited across three research sites: the University of Colorado Boulder (CU; $n = 193$), the University of Missouri (MU; $n = 258$), and the University of Chicago (UC; $n = 34$). Participants were recruited from introductory psychology courses at each site for a study investigating individual differences in cognitive abilities. Recruitment materials specified that participation would involve testing in two sessions separated by approximately 1 week. Participants received either course credit or monetary compensation for completing the testing battery, which took approximately 5 hr. Participants ranged in age from 18 to 42 years ($M = 19.75$, $SD = 2.21$; five participants did not indicate their age). Of the 477 participants who provided demographic information, 31 self-identified as Hispanic, 7 as Native American, 32 as Asian, 37 as Black, 2 as Pacific Islander, and 411 as White (numbers do not sum to 477 because 37 individuals endorsed multiple ethnicities).[2,3] Fn2-3

Of the 485 individuals included in the sample, 406 completed both sessions of the study, forming the sample on which primary hypotheses were tested. Sample sizes for individual tasks varied for additional reasons such as equipment failure or experimenter error. We also excluded data from a given task a priori if performance was below chance, calculated as $p < .01$ binomial probability that the participant would have obtained that score by chance (0.4% of data). With these requirements in place, we had usable data from a minimum of 94.6% of participants for each task. All available data were used in analyses, which allowed for missing observations.

Materials and Measures

Executive function measures. A total of nine computerized EF measures—three tasks tapping each of three EF abilities: response inhibition, working memory updating, and task-shifting—were administered during the first session. These measures form a battery that has been extensively studied by Friedman and Miyake (e.g., Friedman et al., 2006, 2008; Miyake et al., 2000). The particular versions used here were previously used in a large study of young adult twins (Friedman et al., 2014), which found that they show the same factor structure observed here and in other studies (see Miyake & Friedman, 2012). The tasks are described briefly here, but full methodological information can be found in Supplemental Materials.

Response inhibition 1: Antisaccade (adapted from Roberts, Hager, & Heron, 1994). Each trial of this task began with a centrally presented fixation cross that was replaced with an initial cue to the left or right of fixation, after which a numeric target appeared for 150 ms before being masked. Participants' task was to report the target number. In an initial prosaccade block, the

[2] Results do not change when Black participants are omitted from analyses.

[3] Sixty-four participants at CU and 116 at MU completed the bias tasks in the second session while event-related brain potentials (ERPs) were recorded. Those data are not of interest here and will not be reported. We tested whether inclusion of data from individuals who had ERPs recorded in Session 2 affected the results. Across the 12 racial bias constructs assessed (RT bias, accuracy bias, PDP auto, and PDP control for all three tasks), there was only one marginally significant effect ($p = .084$; M p-level $= .64$, *Median* p-level $= .71$). ERP participants had marginally less RT bias on the FPST ($M = -2.45$ ms) than non-ERP participants ($M = 0.48$ ms). We also examined whether ERP participants differed on any construct assessed in the first session (i.e., EF scores and questionnaire factor means). Across the 13 measures there was one significant difference ($p = .015$) and two marginal differences ($ps = .063$ and $.074$; M p-level $= .428$, *Median* p-level $= .310$). ERP participants had significantly higher antisaccade scores ($M = 64.06$) than non-ERP participants ($M = 60.76$), marginally higher stop-signal scores ($M = 250.57$) than non-ERP participants ($M = 244.94$), and marginally lower internal motivation factor scores ($M = -0.11$) than non-ERP participants ($M = 0.59$). ERP and non-ERP participants also showed measurement invariance in their EF models (nonsignificant difference when intercepts and factor loadings were constrained to be equal; $\Delta\chi^2(24) = 35.83$, $p = .057$), which was not surprising given that EF testing in Session 1 was the same for both ERP and non-ERP participants. Most importantly, none of the conclusions about relations between EFs and bias are affected by the inclusion of the ERP subjects as significance levels and directions of effects did not change when the bias task data were removed for the ERP participants.

target always appeared on the same side as the cue to build a prepotency to orient to this stimulus. This was followed by three antisaccade blocks in which the target appeared on the side opposite from the cue. The dependent measure was the averaged proportion of correct responses across the antisaccade blocks.

Response inhibition 2: Stop signal (van den Wildenberg et al., 2006). This task similarly consisted of one block designed to promote a prepotent response followed by three test blocks. On each trial, participants focused on a fixation point until a green arrow appeared that pointed to the right or left. Participants were instructed to press one of two buttons on the keyboard corresponding to the right or left arrows as quickly and accurately as possible. Participants completed an initial block of all-go trials, in which they simply responded according to the arrow direction. After these initial trials, the stop signal was introduced, with the green arrow changing to red on 25% of trials, indicating that participants should inhibit their responses. The onset of this signal was adjusted on a trial-by-trial basis until participants were able to inhibit about 50% of responses. The dependent measure was the *stop signal reaction time* (SSRT; Logan, 1994), which estimates the amount of time required to stop an already-initiated response. The SSRT was calculated as the difference between the median RT on go trials (which estimates the time when a response *would have occurred* in the absence of the stop signal) and the stop signal delay value (i.e., the average stop signal onset across blocks). Larger SSRT values indicate that a participant needed more warning to avoid responding on stop trials (i.e., poorer inhibitory control).

Response inhibition 3: Stroop (Stroop, 1935). This task consisted of three types of trials: (a) neutral trials with strings of three to five asterisks printed in red, blue, or green; (b) congruent trials in which color words were printed in the matching color (e.g., "RED" displayed in red); and (c) incongruent trials in which the word and color never matched (e.g., "RED" displayed in blue). Participants were asked to name the font color aloud. Relative to the neutral trials, congruent trials often show a pattern of facilitation while incongruent trials show a pattern of interference. The overall Stroop effect was calculated as the difference in mean response times between incongruent and neutral trials.

Updating 1: Keep track (adapted from Yntema, 1963). Participants kept track of a series of exemplars belonging to six different categories. Each trial began with a list of two to five target categories (relatives, countries, colors, animals, metals, and distances) shown at the bottom of the screen. Then, a stream of 15 to 25 exemplar words from the various categories appeared in the center of the screen. Participants were asked to verbally recall the most recent exemplar from each target category at the end of the trial. The dependent measure was the proportion of correct responses across all trials.

Updating 2: Letter memory (adapted from Morris & Jones, 1990). As a stream of consonants appeared sequentially on the screen, participants had to rehearse aloud the last four letters seen (including the current letter), in the correct order. Letters were accumulated until the fourth letter was reached, after which the fifth letter back was dropped (i.e., "L," "L-S," "L-S-K," "L-S-K-D," "S-K-D-H," etc.). After 9, 11, or 13 letters had appeared (series length was unpredictable), participants had to report the final 4 letters in the correct order. The dependent measure was the accuracy of the strings repeated after each new letter was presented, with one point given for each correctly reported set.

Updating 3: Spatial *n*-back (Friedman et al., 2008). On each trial a box flashed in one of 12 locations on the screen, and participants reported whether it occurred in the same location as a prior flash. Participants completed both a 2-back and 3-back condition. The dependent measure was the averaged proportion of correct responses across the 2- and 3-back conditions. Omissions were counted as incorrect responses.

Shifting 1: Color–shape (Miyake, Emerson, Padilla, & Ahn, 2004). Participants categorized circles and triangles, presented in either red or green, as quickly and accurately as possible. On each trial, the target was preceded by a cue that indicated whether participants should use the dimension of color ("C") or shape ("S"). The dependent measure was the switch cost: the difference between the average RT for correct switch trials (a trial in which the cue does not match the cue from the previous trial) and correct repeat trials (a trial in which the current cue matches the previous cue).

Shifting 2: Category switch (adapted from Mayr & Kliegl, 2000). This task used a similar structure as the color–shape task but presented words that could be classified both as describing living or nonliving and things smaller or larger than a soccer ball (e.g., *alligator*, *coat*, *knob*, *lion*). A cue (either a heart or crossed arrows) appeared before the word to indicate which dimension was relevant for the current trial. The dependent measure was the switch cost: the difference between average RT for correct switch trials and correct repeat trials.

Shifting 3: Number–letter (adapted from Rogers & Monsell, 1995). On each trial a number-letter or letter-number pair was presented in one quadrant of a square. If the set appeared in the top half of the square, participants were instructed to categorize the *number* as odd or even. If the set appeared in the bottom half, participants were instructed to categorize the *letter* as a consonant or vowel. After initial blocks in which stimuli appeared exclusively in the top half of the square, then exclusively in the bottom half, a block of predictable switches occurred in which the pair of characters was presented in a clockwise pattern. Participants then performed random-switch blocks in which the stimulus's location was randomly determined on each trial. The dependent measure was the switch cost in these random-switch blocks: the difference between average RTs for correct switch trials and correct repeat trials.

Implicit measures of racial bias. Three computerized measures of racial bias, all designed to assess the same basic aspect of racial stereotypes of African Americans (the association between young African American men and danger), were administered in the second session.

Weapon Identification Task (WIT). The WIT (Payne, 2001) required participants to classify objects as either guns or tools. The task included a practice block of 30 trials, followed by a test block of 384 experimental trials.[4] On each trial, a visual pattern mask (a Fn4

[4] There were an additional 32 trials each in which a Black or White prime was shown but no target gun/tool stimulus followed. These trials were inserted for the purpose of ERP analyses (to allow quantification of ERP responses to primes into the period where the target stimulus would normally appear). These prime-only trials were randomly intermixed among the prime-target trials. Participants were instructed to omit responses on any trial lacking a gun or tool. Because there are no behavioral responses on these trials, they are not considered in the present analyses.

scrambled black and white pattern; 500 ms duration) preceded a briefly presented picture of a White or Black male face (i.e., prime, 200 ms duration), which was followed immediately by presentation of either a gun or tool (i.e., target, 200 ms duration). The target was then hidden by a second visual mask (300 ms duration). Participants were told that the faces served only as a cue that a target object was about to appear, and that they should classify the objects as guns or tools as quickly and accurately as possible by pressing one of two buttons. Responses made following a 500 ms response deadline elicited a "Too Slow!" message to encourage faster responses. Trials were separated by a 1,000-ms intertrial interval. Previous studies have consistently shown performance indicative of racial bias on this task in that participants more frequently press the "gun" key following Black face primes compared with White face primes (e.g., see Amodio et al., 2008, 2004; Bartholow et al., 2012; Payne, 2001, 2005).

The race primes consisted of eight pictures each of Black and White males displaying neutral facial expressions. They were shown in color and cropped to show only internal face features. Target stimuli were grayscale images of four handguns and four tools (a drill, two wrenches, and pliers) from Payne (2005). The dependent measure of interest for the WIT was bias in response accuracy, calculated as the difference in accuracy on stereotype-congruent trials (Black–gun and White–tool trials) and incongruent trials (Black–tool and White–gun trials).

First-Person Shooter Task (FPST). The FPST (Correll et al., 2002) requires participants to make speeded decisions to either shoot or not shoot images of armed and unarmed Black and White men. The task began with a practice block of 16 trials followed by a test block of 300 trials. On each trial, participants first saw a variable number of background images (one to four images, 500 ms–800 ms duration). The final background image was replaced with an image of a target individual in that same background, leaving the impression that the target popped up in the scene. The target was presented for 590 ms. The target was holding either one of several handguns or one of several innocuous objects (e.g., cell phone, wallet, soda can). Participants were instructed to press a button labeled "Shoot" (for armed targets) or a button labeled "Don't Shoot" (for unarmed targets) as quickly as possible. If a response was made within the 590-ms response window (i.e., while the target was displayed), participants received the following feedback and points added to a running score (presented for 500 ms): (a) correct decision to shoot—"Good shot," 10 points; (b) incorrect decision to shoot—"You shot a good guy!!" 40 point penalty; (c) Correct decision not to shoot—"Wise choice," 5 points; and (d) incorrect decision not to shoot—"YOU'RE DEAD!!" 20 point penalty. Timeouts were penalized with a 50-point deduction. Previous studies have shown patterns of bias on this task such that participants more frequently shoot Black than White targets (see Correll et al., 2002, 2006). Performance bias for this task was calculated as the differences in accuracy on stereotype congruent trials (armed Black and unarmed White) and incongruent trials (unarmed Black and armed White trials).

Implicit Association Test (IAT). On each trial of this task (adapted from Greenwald et al., 1998), participants saw a single stimulus: either a picture of a male face (Black or White) or a word connoting either safety or danger. Participants were asked to categorize each stimulus as quickly and accurately as possible by pressing one of two buttons. Importantly, the buttons used to categorize the stimuli changed during the course of the task. Below, we report the response categories assigned to the right-hand button (the alternate categories were assigned to the left). Stimuli remained on the screen until participants responded (a correct response was not required and no error feedback was given). Seven blocks were completed in the following order: (1) classifying faces in terms of race (24 trials, right button = Black); (2) classifying words as representing safety or danger (24 trials, right button = danger); (3) classifying both faces and words with a stereotype-congruent response mapping (48 trials, right button = Black or danger); (4) a second block identical to Block 3 (96 trials); (5) classifying only faces, but with the original response mapping reversed (24 trials, right button = White); (6) classifying both faces and words with a stereotype-incongruent response mapping (48 trials, right button = White or danger); and (7) a second block identical to Block 6 (96 trials). During the dual categorization blocks (3, 4, 6, and 7), stimulus type (face vs. word) alternated. Previous studies (e.g., Amodio & Devine, 2006; Cunningham et al., 2004; Glaser & Knowles, 2008; Hugenberg & Bodenhausen, 2004; Richeson & Shelton, 2003) have shown that participants tend to categorize stimuli in the congruent response mapping blocks (i.e., Black–danger; White–safety) more quickly than in the incongruent response mapping blocks (i.e., Black–safety; White–danger), indicating racial bias.

Target stimuli consisted of 12 pictures each of Black and White men, matched for attractiveness based on pilot data. Safety words were *nice, kind, friend, trust, peaceful, happiness, protected, secure, harmony, unity,* and *caring.* Danger words were *violence, aggression, mean, brutal, nasty, attacking, knife, enemy, fight, harmful, cruel,* and *fear*.

Racial bias in RTs was calculated as a *D*-score (as recommended by Greenwald, Nosek, & Banaji, 2003). This involved calculating two scores: (a) one value derived from subtracting the average RT in the first set of compatible and incompatible blocks (Blocks 3 and 6) divided by the pooled standard deviation of these two blocks; and (b) a second value calculated in the same manner using the second set of compatible and incompatible blocks (Blocks 4 and 7). These were then averaged to obtain the final *D*-score. Trials with RTs < 200 ms or > 10,000 ms were dropped, as were participants who had RTs < 300 ms on more than 10% of trials. Following Greenwald, Nosek, and Banaji (2003), response latencies on incorrect trials were replaced with the block mean + 600 ms.

Self-report questionnaires. Participants completed the following questionnaires during Session 1: (a) Attitudes Toward Blacks (Brigham, 1993); (b) Motivation to Control Prejudiced Reactions scale (Dunton & Fazio, 1997); and (c) Internal and External Motivation to Respond without Prejudice (Plant & Devine, 1998). Participants also completed feeling thermometers to assess racial attitudes, rating how they feel toward several social groups (Gay men, Hispanics, African Americans, Lesbians, Asian Americans, and White Americans) on a scale from 0 (*very coolly*) to 100 (*very warmly*). Participants' personal stereotype content was measured using semantic differential line scales assessing the extent to which they perceived African Americans and Caucasians as possessing the following characteristics: aggression (anchored with *not aggressive* and *aggressive*), violence (anchored with *not violent* and *violent*), and dangerousness (anchored with *not dangerous* and *dangerous*). These three items were completed separately for African Americans and Caucasians. Cultural stereotype perceptions were measured by having participants respond to the same items but instructing them: "instead of telling us what *you* believe, we want you to tell us how you think *most Americans* would answer these questions." Bias variables were created for the feeling thermometer, personal stereotype, and cultural stereotype questions by subtracting participants' ratings for the separate White and Black items such that higher scores represented more pro-White/anti-Black bias.

Finally, we administered an indirect measure of race bias, a student group funding allocation task (Correll, Park, & Smith, 2008). Participants were told that data was being collected to determine which student groups should receive funding in light of an $11 million reduction in the university budget. They were then given a list of 30 student groups, and asked to rank order them in terms of program importance (1 as the most important, 30 as the least). Embedded in this list were four directly related to African Americans or racial minorities more generally: the African American student association, the summer minority access to research training program, the minority arts and science program, and the prospective freshman minority visit program. The average rank of these four programs was subtracted from the average rank of four nonminority "neutral" student groups (i.e., the university community service center, the community health initiative, the fitness for life program, and the undergraduate research symposium) to create a measure of anti-Black racial bias in funding allocations.

Data Scoring and Analysis

Questionnaire scoring. As highlighted throughout the work, a primary strength of our current approach is our ability to use multiple related measures to extract latent variables, which represent more accurate estimates of our conceptual variables. We applied this method to our self-report measures, entering all race-related items from the explicit measures into an exploratory factor analysis. To do this, the individual questionnaire items and the various bias scores (55 total items) were used to create a set of factor scores that more accurately capture the various constructs measured. Using principal axis factoring and promax rotation, we extracted four primary factors based on the scree plot and an eigenvalue cutoff of 2.0. Together these four factors explained 38% of the total variance. After examining the individual item loadings from the pattern matrix (see Appendix Table A1), we labeled these factors: (a) Personal Racial Attitudes (eigenvalue = 10.60, 19% of variance explained); (b) Internal Motivation to Control Prejudiced Responding (eigenvalue = 5.35, 10% of variance explained); (c) External Motivation to Control Prejudiced Responding (eigenvalue = 2.77, 5% of variance explained); and (4) Perceived Cultural Stereotypes (eigenvalue = 2.21, 4% of variance explained). To confirm our conceptual labels of these factors, we also examined factor correlations with the individual scales (scored as outlined by their original authors). As can be seen in Appendix Table A2, these correlations provide further evidence that the factors do indeed capture the conceptual variables described above. Extracted factors scores were used in all subsequent analyses.

Though not of focal interest, there are some noteworthy patterns in the data from this factor analysis (see Appendix Table A1). Although the various items we included were created to measure very specific constructs, many of them load onto multiple factors, contributing substantial variance to personal attitudes, internal motivation to control prejudiced responding, and external motivation to control prejudiced responding. Also of note is the fact that despite separately measuring both personal and cultural stereotype endorsement (using the same scale and question formats), only the cultural stereotypes formed their own factor. Personal stereotypes were primarily captured by the Personal Attitude factor.

PDP scoring. For each of the three implicit measures of racial bias, we performed a PDP analysis based on error rates. This analysis attempts to separately estimate the extent to which participants' responses were determined by control-related processes and "automatic" tendencies to indicate the presence of a threat. Following prior research (Payne, 2001, 2005), these control and automaticity estimates were computed separately for trials involving White and Black stimuli in the WIT and FPST. The control formulas represent the extent to which participants perform as intended given the goals of the task, and were calculated by subtracting the probability of an error on stereotype incongruent trials (e.g., a tool following a Black face prime in the WIT) from the probability of a correct response on a stereotype congruent trial (e.g., a gun following a Black face prime in the WIT). The automaticity formulas represent the extent to which stereotypes promote a "gun" response when control fails, and were calculated from the probability of an error on a tool trial, divided by one minus the control estimate. Past research has shown a greater contribution of automatic processes on trials with Black rather than White stimuli, yet no difference in control on Black and White trials (see Payne, 2001, 2005). Moreover, to the degree that the control estimates reflect ability to follow task requirements (e.g., identify a gun or tool while making no response to the primes in the WIT), PDP control should reflect a single underlying construct (Amodio et al., 2008). For each task, we therefore computed two primary indices on interest: (a) racial *bias* in the contribution of automatic processes to participants' responses ([the automatic tendency to indicate danger induced by Black stimuli] − [the automatic tendency to indicate danger induced by White stimuli]); and (b) *mean* level of control (i.e., the average of control estimates induced by both White and Black stimuli).

Because of its more complicated task structure, PDP calculations were somewhat different for the IAT. For this task, we computed four separate control and automaticity estimates, one each for Black and White faces, and one each for danger and safety words. Each of these four computations followed the basic pattern outlined above, with control calculated as the probability of an incorrect response during a stereotype-incongruent block (e.g., a miscategorization of a Black face when Black and safety shared a response key) subtracted from the probability of a correct response during a stereotype-congruent block (e.g., a correct categorization of a Black face when Black and danger shared a response key). We then created two interim automaticity bias and mean control variables, one for face trials and another for word trials. These were averaged to obtain our final estimates of automaticity bias (representing the average amount of racial bias expressed on face and word trials) and mean control (representing the mean level of control across all trial types). Although PDP computations have been performed on the IAT (Huntsinger, Sinclair, & Clore, 2009; Stewart, von Hippel, & Radvansky, 2009), they are done so less frequently than with other measures of implicit bias. As a consequence, they might be viewed as more exploratory.

Data trimming and transformation. Following our previous studies (e.g., Friedman et al., 2008), we applied appropriate trimming and transformation to the EF data to improve the distributions and reduce the influence of outliers. First, for RT measures (shifting and Stroop), response latencies were trimmed according to Wilcox and Keselman (2003; Equation 3). After this within-subject trimming, we also conducted between-subjects trimming: For each measure, scores greater than 3 *SD* from the group mean were replaced with a value that was equal to 3 *SD* above or below the group mean, as appropriate. This replacement technique aimed to preserve each participant's rank ordering, while preventing extreme outliers from unduly influencing correlations or model parameters. The short response windows used in the WIT and FPST resulted in a smaller range of response latencies, obviating the need for trimming. As noted earlier, we followed Greenwald et al.'s (2003) recommendation for the treatment of fast and slow response latencies in the IAT rather than employing the Wilcox and Keselman trimming.

There were two bias-related measures that were not normally distributed: factor scores for the first questionnaire factor (Personal Attitude), and PDP control estimates for the IAT task. Normality was improved with natural log and arcsine transformations, respectively.

Analyses of response latencies in the WIT and FPST tasks were performed on log-transformed values, although data are reported in ms for ease of interpretation. After these trimming and transformation procedures were applied, all measures used in the structural equation models had skewness and kurtosis values between -1 and 1. For ease of interpretation, we reversed the directionality of EF RT variables (stop-signal, Stroop, and switch tasks) in all analyses so higher numbers indicate better performance.

Model estimation. Structural equation models were estimated using Mplus 6.12 to 7.2 (Muthén & Muthén, 1998–2012), which includes participants with missing data for one or more measures. (Models used all available data, which is why the *N*s noted in the tables sometimes differed.) Because the χ^2 is sensitive to sample size, we also used confirmatory fit index (CFI) $<$ .95 and root-mean-square error of approximation (RMSEA) $<$ .06 as indicators of good fit (Hu & Bentler, 1998). Statistical significance of parameters was established with χ^2 difference ($\Delta\chi^2$) tests.

Models that included interactions between latent variables were estimated with TYPE = RANDOM and numerical integration, which provided full-information maximum likelihood estimates of the interactions (Klein & Moosbrugger, 2000); Mplus does not provide fit statistics for these models but does provide log-likelihood and scaling values that can be used for nested model comparisons (Satorra & Bentler, 2001). In cases when the calculation of the scaled $\Delta\chi^2$ from these comparisons resulted in a negative value (an improper value that can indicate that the constrained model is highly incorrect; Satorra & Bentler, 2010), we used Wald statistics to test significance. To assess overall fit for these models, the Mplus Web site FAQ on latent variable interactions (http://www.statmodel.com/faq.shtml) suggests ascertaining that the model without interactions fits the data well and then assessing whether added interactions are significant; this strategy is appropriate because chi-squares for models without interactions are not sensitive to leaving out the interaction (Mooijaart & Satorra, 2009) due to their noninclusion of higher-order moments (beyond second-order variances and covariances). Hence, for models with interactions, we present fit statistics for the same models without interactions. Standardized interaction terms were calculated according to the method described in Wen, Marsh, and Hau (2010).

Procedure

Across two testing sessions, roughly 12 days apart, participants completed nine EF tasks, three implicit racial bias tasks, several race-related questionnaires, and demographic measures. All EF tasks and questionnaires were completed during the first 3-hr testing session; the three implicit racial bias tasks were completed during the second 2-hr testing session.[5] Fn5

Because this study was designed expressly to examine individual differences, all participants completed both sessions' tasks in a single, fixed order to minimize order effects. The first session tasks were administered in the following order: stop signal, demographic questions, spatial 2-back, category switch, Stroop, funding allocation measure of racial bias, 3–5 min break, keep track, color/shape, letter memory, Motivation to Control Prejudice Responses scale (Dunton & Fazio, 1997), Internal and External Motivation to Control Prejudice scale (Plant & Devine, 1998), feeling thermometer, 3–5 min break, antisaccade, number/letter, spatial 3-back, stereotype content measure, and Attitudes toward Blacks. During the second testing session, the three implicit measures of racial bias were administered in the following order: WIT, IAT, 3–5 min break, and FPST.

All computerized tasks were administered on Macintosh computers (iMacs running OS X Leopard) and programmed in PsyScope, except for stop signal (an executable program run from a Windows XP partition on the iMacs). Responses were made using ms-accurate button boxes (IoLabs, United Kingdom) for all tasks except the Stroop, for which the timing of verbal responses was recorded with a headset microphone attached to this button box, and the stop-signal, which used a ms-accurate keyboard (Empirisoft, Inc, New York). For all tasks except the implicit measures of race bias, stimuli and trial order were presented in a single, fixed pseudorandom order to all participants. For the implicit measures of bias, the computer selected a new random order for each participant.

Results and Discussion

Before evaluating the theoretical issues of primary interest concerning relations between EF and implicit bias, we first examine basic descriptive aspects of our EF and implicit bias tasks. We specifically confirm the nested factor EF structure through a latent

[5] Participants also performed De Houwer's (2003) Extrinsic Affective Simon Task. However, in line with a recent critique offered by De Houwer himself (De Houwer & De Bruycker, 2007), this task showed poor split-half reliability and did not correlate meaningfully with the other measures of bias in this study. Because of its problematic status in the literature, poor psychometric performance, and the impossibility (based on low intercorrelations) of employing it in latent variables analyses, we excluded it from the final analysis. It is crucial to note that this task was always administered at the end of the second session (after all of the other tasks) and therefore could not influence performance on them.

variable analysis, expecting Common EF, Updating-specific, and Shifting-specific factors, as described by Miyake and Friedman (2012). Next we confirm the efficacy of our implicit bias tasks, expecting mean levels of performance indicative of racial bias within each task. We then consider the important issue of interrelations among the bias measures, which we evaluate in terms of *performance bias* (i.e., being faster and more accurate on stereotype congruent than incongruent trials) as well as the contribution of automatic and controlled processes, as reflected in process dissociation estimates computed within the PDP framework (Jacoby, 1991).

Preliminary Analyses

Executive function model. Means, standard deviations, and reliabilities for each of the EF tasks are shown in Table 1, and their correlations are in Table 2. Figure 1 shows the nested factors EF model described by Miyake and Friedman (2012). These analyses were based on the sample of 484 participants who had usable EF data. Consistent with previous work (e.g., Friedman et al., 2011), we find Common EF, Updating-specific, and Shifting-specific factors. Inspection of the fit statistics and modification indices suggested a residual correlation between antisaccade and spatial *n*-back. After adding this post hoc correlation, the model fit well, $\chi^2(20) = 32.01$, $p = .043$, CFI = 0.980, RMSEA = 0.035. We use this EF model in all subsequent analyses. As with past research (Friedman et al., 2011), we did not find evidence for an Inhibition-specific factor within this model. Specifically, the three inhibition tasks showed nonsignificant loadings on such a factor, and the residuals for the inhibition tasks did not significantly positively correlate with each other.[6]

T1 T2, F1 Fn6

Implicit bias tasks. Racial bias in accuracy was of primary interest in WIT and FPST due to the short response deadline used in both tasks, whereas racial bias in RT was of primary interest in the IAT due to its lack of response deadline. We nevertheless calculated bias in both types of responses for all measures. Mean performance bias, *SD*s, and reliabilities for each of the implicit bias tasks are shown in Table 1. Also included are tests of mean levels of racial bias, all coded such that higher numbers indicate that the association between Blacks and danger is stronger than the association between Whites and danger. Because these analyses use data from Session 2, they are based on the 406 participants who completed Session 2. As expected, racial bias (indicating a stronger association between Blacks and danger) was present across all tasks (although not significantly so in RT on FPST, likely due to the tight response deadline).[7] Mean accuracy and RTs within each condition for each task are shown in Appendix Table A3.

Fn7

PDP control and automaticity estimates for each task are shown in Table 3. The current data replicate past findings of relatively greater automaticity on trials involving Blacks than Whites (e.g., Payne, 2001, 2005). Although prior work typically does not find racial differences in the PDP control estimates, the power afforded by our large sample revealed significant race differences in PDP control in all three tasks, albeit in differing directions and of relatively small magnitude. Given the inconsistent and small effects (and lack of precedent for such findings) we caution overinterpretation of this result. Correspondingly, we follow convention throughout the remainder of the analyses and focus on *mean* PDP estimates of control and *bias* in PDP estimates of automaticity (i.e., PDP auto on Black trials—PDP auto on White trials; see Payne, 2005). Note that although error rates were lower in the IAT than WIT and FPST (see Table A3), the IAT PDP control estimate had high reliability (see Table 3). Reliability for PDP auto was more modest for the IAT, but exceeded that of the FPST. Moreover, the large number of trials we administered in each task resulted in higher absolute numbers of errors than is often obtained with these tasks (Table A3).

T3

We next examined bivariate correlations between the racial bias measures (see Table 4), which are particularly relevant for ascertaining whether the different implicit bias measures assess the same underlying constructs and, hence, whether they can be used to form latent variables. These correlations show that WIT and FPST are most closely related (despite the very low internal reliability estimates for FPST). At the level of overall performance, accuracy bias for WIT and FPST were significantly correlated. The *D*-score measure that is most traditional for the IAT did not significantly correlate with the accuracy bias scores for the other two tasks (although there was a small correlation between WIT accuracy bias and IAT accuracy bias). The pattern was similar for the PDP auto measures, with only WIT and FPST correlating. There were no significant relationships among the three RT measures. Thus, despite the fact that all three implicit measures were designed to assess the same stereotype content (i.e., mental associations between race and weapons/danger), it appears that there are some very real differences in the constructs they measure and/or the processes they reflect. T4

Although the tasks appear to dissociate in performance bias and racial bias in PDP auto, very different conclusions are evident when examining the PDP control estimates. Here, all three tasks do appear to be measuring similar constructs, evident both in terms of the magnitude of the relation and the number of tasks correlating with each other. Mean level of control was correlated across all three tasks, suggesting substantial overlap in the control assessed by each task. It is important to note, however, that although IAT PDP control was correlated with the control estimates from the other two tasks (.37 and .29) those correlations were substantially lower than the correlation between control estimates for WIT and FPST (.61). We will return to this issue in more detail under discussion of H3.

Together, these results suggest the following conclusions. Our tasks were effective in assessing implicit racial stereotyping; replicating scores of prior studies, our participants associated Blacks

[6] There was a significant negative correlation between the residuals for antisaccade and Stroop, $r = -.68$, $\chi^2(1) = 16.27$, $p < .001$. Because the correlation was negative and was also inconsistent with our findings in another sample, in which the only significant residual correlation among the response inhibition tasks was a negative correlation between stop-signal and Stroop (Friedman et al., 2011), we did not consider it as evidence for an Inhibition-specific factor and did not include it in the final models.

[7] The magnitude of behavioral bias in the FPST was somewhat smaller than we typically see (eta-squared = .06). This may be due to the extremely short response window (590 ms vs. 630 ms, which is typically used) or, more likely, to the fact that participants performed 300 trials, rather than the standard 100 trials. Indeed, previous work has shown that practice with this task can reduce the magnitude of bias (Correll et al., 2007).

Table 1
Descriptive Statistics

Dependent measure	Mean (*SD*)	*N*	Reliability	Tests of Racial Bias: *t*-value	Tests of Racial Bias: *p*-value
Executive function tasks[a]					
Response inhibition					
Antisaccade	61.98 (14.36)	482	.80		
Stop signal	247 ms (32)	458	.41		
Stroop	141 ms (70)	477	.94		
Updating					
Keep track	73.32 (8.60)	483	.60		
Letter memory	75.12 (13.49)	482	.91		
Spatial *n*-back	80.73 (6.80)	470	.84		
Shifting					
Number–letter	243 ms (156)	480	.92		
Color–shape	241 ms (160)	479	.88		
Category switch	156 ms (119)	481	.92		
Implicit measures[b]					
Weapons Identification Task					
Accuracy performance bias	0.17 (0.19)	401	.74	17.99	<.001
Reaction time performance bias	13 ms (23)	401	.46	11.25	<.001
First Person Shooter Task					
Accuracy performance bias	0.03 (0.11)	391	.09	5.15	<.001
Reaction time performance bias	−0.80 ms (16)	391	.07	1.22	.22
Implicit Association Test					
Accuracy performance bias	0.03 (0.04)	405	.67	16.21	<.001
Reaction time (*D*-score) bias	0.47 (0.27)	402	.87	34.89	<.001
Explicit measures					
Attitudes toward Blacks	2.37 (0.84)	466	.86		
Internal motivation to control prejudice	6.97 (1.75)	481	.85		
External motivation to control prejudice	4.99 (1.84)	481	.83		
Motivation to control prejudiced Reactions	4.22 (0.79)	482	.81		
Feeling thermometer bias	10.46 (21.28)	480	—	10.77	<.001
Personal stereotype bias	0.29 (1.65)	483	.80	3.85	<.001
Perceived cultural stereotype bias	3.88 (2.39)	483	.91	35.63	<.001
Funding allocation bias	5.93 (6.68)	481	—	19.47	<.001

Note. Reliability assessed with odd–even reliability according to Cohen and Cohen (1984). Test of racial bias is a single-sample *t*-test comparing the variable's mean to zero.
[a] For EF tasks that are scored based on accuracy, higher values indicate better EF but for EF tasks that are scored based on reaction time, higher values indicate poorer EF. For model estimation, EF reaction time variables were reversed so that higher numbers indicated better EF for all EF variables. [b] Descriptives shown are for the untransformed variables. Models used transformed data to improve normality.

more strongly with threat and danger than they did Whites. However, while all tasks were designed to assess these same mental associations, performance across tasks was only weakly related. This was especially true of the IAT, as the strongest and most consistent relations were between WIT and FPST. Across the different types of information provided by these tasks (i.e., accuracy performance bias, RT performance bias, PDP auto, PDP control), they covary most strongly in terms of the level of control they elicit.

Effect of Automatic Processes and Executive Functions on Implicit Bias (H1 and H2)

We evaluated our primary theoretical questions regarding the relation between EF and implicit bias through structural equation modeling using latent variables for our constructs of interest. The dissociation of IAT from WIT and FPST (see Table 4) precluded formation of a single latent implicit bias variable from all three tasks. In the sections that follow, we present two separate instantiations of each analysis, first using a latent implicit bias variable constructed from performance on the WIT and FPST tasks, then conducting parallel analyses using IAT performance. Figure 2 shows the latent accuracy performance bias, PDP auto, and PDP control variables created from the WIT and FPST, estimated separately. F2

WIT and FPST. Our first model was designed to examine the degree to which implicit racial bias is a function of (a) EF, (b) automatic processes as reflected in PDP auto estimates, and (c) the interaction of EF and PDP auto. As described in the earlier Model Estimation section, we first estimated a model without interactions, which resulted in good fit, $\chi^2(51) = 71.95$, $p = .028$, CFI = 0.989, RMSEA = 0.029. Note that we allowed the residual variances for the bias measures from the same tasks

Table 2
Correlations Among EF Measures

Task	1	2	3	4	5	6	7	8	9
1. Antisaccade	—								
2. Stop-signal	.16*	—							
3. Stroop	.18*	.09	—						
4. Keep-track	.19*	.03	.17*	—					
5. Letter memory	.32*	.02	.30*	.37*	—				
6. Spatial *n*-back	.39*	.04	.14*	.31*	.38*	—			
7. Number-letter	.18*	.11*	.23*	.07	.12*	.11*	—		
8. Color-shape	.07	.00	.12*	.02	.06	.06	.37*	—	
9. Category-switch	.25*	.10*	.16*	.08	.12*	.11*	.50*	.41*	—

Note. Correlations are maximum likelihood estimates (from Mplus) based on all data, adjusted for missing observations. Total $N = 484$. All data were scored so that higher numbers indicated better performance.
* $p < .05$.

to correlate to account for task-specific method covariance (i.e., the variance in WIT performance bias explained by the performance bias latent variable was allowed to correlate with the variance in WIT PDP auto not explained by the PDP auto latent variable, and the same for FPST). We then added interactions between PDP auto and EF to test H2. We started with preliminary models that included all three aspects of EF from Figure 1 (Common EF, Updating-specific, and Shifting-specific) as well as interactions of all three with PDP auto. There were never any significant interactions of the Updating-specific or Shifting-specific factors with PDP auto in the WIT/FPST analyses. For the sake of clarity, therefore, all subsequent models of implicit bias in WIT and FPST include only the Common EF x
F3 PDP Auto interaction. This model, shown in Figure 3, confirmed our predictions. Specifically, we found that both racial bias in automatic processes and Common EF predict performance bias. As specified in H1, racial bias increased as a function of relatively greater automatic processes on Black than White trials, Wald test (1) = 65.46, $p < .001$, but decreased as a function of Common EF, Wald test (1) = 21.65, $p < .001$.

This model also considered that the relationship between PDP auto and bias might be moderated by EF (e.g., Payne, 2005). To the degree that EF is needed to exert control over automatically activated racial associations, we would expect the relationship between automatic processes and performance bias to be weaker for those with greater EF (H2). This predicted Common EF × PDP Auto interaction was significant, Wald test
F4 (1) = 45.29, $p < .001$. The interaction is depicted in Figure 4, with low and high values of Common EF estimated at −/+ 1 *SD*. As can be seen, the positive relation between PDP auto and implicit bias occurs for both those low (simple slope $b = 1.25$, $SE = 0.14$, $z = 8.84$, $p < .001$) and high (simple slope $b = 0.63$, $SE = 0.11$, $z = 5.95$, $p < .001$) in Common EF but is weaker for those high in Common EF, as hypothesized. As a result, at high levels of PDP auto, implicit bias is greater for those lower in Common EF (simple slope $b = -0.54$, $SE = 0.08$, $z = -6.84$, $p < .001$). By contrast, when PDP auto is low, implicit bias is not significantly different for those low and high in Common EF (simple slope $b = 0.08$, $SE = 0.05$, $z = 1.49$, $p = .137$). Together, these results validate and extend the findings of prior research. As others have documented, participants' responses on implicit bias measures are significantly contaminated by individual differences in EF (Klauer et al., 2010). Those individuals with high levels of EF display lower levels of bias. Although one plausible explanation for this effect is the presence of stronger racial stereotypes among low EF-individuals, the interaction observed between Common EF and PDP auto suggests that more complex processes are at play. For participants relatively low in EF, bias in PDP auto translates more directly into performance bias on the implicit measures. Participants relatively high in EF, however, are better able to prevent these biased associations from influencing their behavior. The pattern of these results also suggests that, when PDP auto is low, EF has virtually no effect. This suggests that, in the absence of a stereotypic impulse, EF does not matter. However, when PDP auto is high, suggestive of a strong stereotypic response tendency, EF ability plays a critical role. Participants who are higher in EF seem better able to avoid behaving in line with their automatic tendencies.

IAT. To evaluate the contribution of EF to performance on the IAT, we estimated an additional structural equation model in which just the IAT RT based *D*-score was regressed on the EF variables (see Figure 5). The model without interactions fit the data F5
well, $\chi^2(32) = 49.25$, $p = .026$, CFI = 0.976, RMSEA = 0.033. Given that an association between IAT performance and shifting has been demonstrated before (Klauer et al., 2010), we present the full model with interactions between all three EF latent variables and PDP auto.

This model revealed two significant predictors of IAT bias. First, replicating the relationship found in the WIT/FPST analysis, relatively greater racial bias in automaticity was associated with greater IAT performance bias, Wald test (1) = 94.79, $p <$

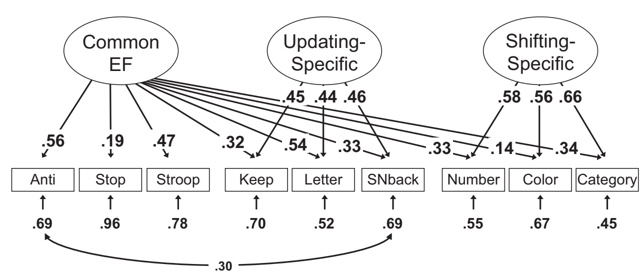

Figure 1. Executive function (EF) model. Ovals indicate latent variables and rectangles indicate observed measures. Numbers on straight arrows are standardized loadings, numbers on curved, double-headed arrows are correlations, and numbers at the end of arrows are residual variances. There is a Common EF latent variable on which all nine executive function tasks load, as well as two "nested" latent variables on which the updating and shifting tasks, respectively, also load. The Common EF variance was isomorphic with the inhibition latent variable, so there was no Inhibition-specific variance. Because the Common EF factor captures the variance common to all three EFs, the Updating-specific and Shifting-specific factors capture the variance that is unique to updating and shifting, respectively. Hence, they are uncorrelated with the Common EF factor and with each other. In all models, EF variables were scored so that higher numbers indicate better performance. Anti = antisaccade, Stop = stop-signal, Letter = letter memory, SNback = spatial *n*-back, Number = number-letter, Color = color-shape, Category = category-switch. Boldface type indicates $p < .05$.

Table 3
Process Dissociation Procedure Parameter Estimates for Implicit Bias Measures

Bias measure	Control	*t*-value	Automaticity	*t*-value
Weapons Identification Task	*.92^{rel}*		*.64^{rel}*	
Black	.48 (.19)		.61 (.14)	
White	.50 (.18)		.45 (.15)	
Racial bias[a]	−.02 (.11)	3.31*	.16 (.17)	19.44*
First Person Shooter Task	*.91^{rel}*		*.14^{rel}*	
Black	.50 (.20)		.56 (.13)	
White	.45 (.17)		.53 (.12)	
Racial bias[a]	.05 (.10)	9.72*	.04 (.11)	6.64*
Implicit Association Test	*.93^{rel}*		*.36^{rel}*	
Black	.91 (.07)		.59 (.19)	
White	.89 (.09)		.59 (.19)	
Danger	.78 (.15)		.58 (.19)	
Safety	.83 (.12)		.68 (.16)	
Racial bias[b]	−.01 (.06)	4.72*	.22 (.19)	22.77*

Note. Standard deviations in parentheses. The *t* values represent tests of the magnitude of bias against zero, showing whether there is differential control or automaticity associated with trials involving Blacks compared to Whites.
[a] Racial Bias = Black estimate–White estimate. [b] Racial Bias = [(Black Estimate − White Estimate) + (Danger Estimate − Safety Estimate)]/2. *Rel* = reliability.
* $p \leq .001$.

.001. Second, consistent with prior demonstrations that the IAT contains a shifting component (Klauer et al., 2010), Shifting-specific ability was negatively associated with IAT bias, $\Delta\chi^2(1) = 21.80$, $p < .001$, indicating that greater Shifting-specific ability was associated with the expression of less racial bias on the IAT. In contrast with the model predicting bias in WIT and FPST (see Figure 3), Common EF was unrelated to IAT bias, $\Delta\chi^2(1) = 0.03$, $p = .861$, and there were no interactions between any aspect of EF and PDP auto (H2), all $\Delta\chi^2(1) < 2.67$, $p > .102$.

Relationship of Executive Functions With PDP Control and Auto (H3)

EFs have been theorized to play a role in PDP control but not auto (Payne, 2005). To test this prediction and to assess which aspects of EF relate to PDP control, we estimated a series of structural equation models in which WIT, FPST, and IAT PDP control and auto values were separately regressed on the EF variables. Standardized path coefficients are shown in Table 5. All of these models fit the data well, T5 all $\chi^2(26) < 43.41$, $p > .017$, CFI > 0.972, RMSEA < 0.038.

Consistent with H3, EFs predicted PDP control but not auto. However, the pattern of relations between EF and PDP control differed across bias tasks. Although bivariate correlations among the PDP control estimates showed significant relations among all three tasks (*r*s ranging from .29 to .61; see Table 4), these correlations obscured important differences in the types of control elicited from the tasks. While control in the WIT and FPST was related to Common EF, control during the IAT was not. Instead, IAT control was related to both Updating-specific (positively) and Shifting-specific (negatively) abilities. We had no a priori reason to expect a relation with the

Table 4
Intercorrelations Among Implicit Measures and Their Associated PDP Estimates

Measure	1	2	3	4	5	6	7	8	9	10	11	12
Accuracy bias												
1. WIT	—											
2. FPST	.18*	—										
3. IAT	.11*	.02	—									
Reaction time bias												
4. WIT	.37*	.13*	−.02	—								
5. FPST	−.01	−.09	.05	.01	—							
6. IAT (*D*-score)	.07	.05	.45*	.06	.01	—						
PDP auto estimate—bias												
7. WIT	.89*	.20*	.05	.40*	.02	.09	—					
8. FPST	.12*	.91*	−.05	.15*	−.03	.02	.16*	—				
9. IAT	.10*	.05	.77*	.00	.06	.48*	.10	.02	—			
PDP control estimate—mean												
10. WIT	−.26*	.05	−.18*	.10	.05	.04	.05	.11*	.02	—		
11. FPST	−.09*	.14*	−.13*	.18*	.06	.01	.08	.22*	.03	.61*	—	
12. IAT	−.03	.10	−.50*	.06	−.01	−.10*	.06	.14*	.00	.37*	.29*	—

Note. Correlations are maximum likelihood estimates (from Mplus) based on all data, adjusted for missing observations. Total $N = 406$. RT = reaction time.
* $p < .05$.

Updating-specific factor, and the relation with Shifting-specific abilities indicated that the ability to follow task instructions during the IAT declined as Shifting-specific abilities increased. Notwithstanding the difficulty of interpreting these relations, they (as well as those in Table 4) indicate that the PDP-derived IAT control scores reflect a very different process than that assessed by the PDP control scores for the WIT and FPST tasks.

EF Moderates the Relation Between Implicit and Explicit Bias (H4)

Our next set of analyses examined the moderating effect of EF on the relation between implicit and explicit measures of racial bias. These measures are often found to correlate only weakly (e.g., Blair, 2001; Cunningham et al., 2001; Fazio & Olson, 2003), which has been attributed to the greater susceptibility of explicit measures to controlled processes (e.g., Koole et al., 2001; Ranganath et al., 2008). If so, EF and implicit bias may interact in predicting explicit bias, with implicit measures more strongly predicting explicit responses for people who have low EF (H4). The bivariate correlations between the implicit bias measures and
T6 the factor scores from the explicit bias measures shown in Table 6 confirm the general expectation of low relations between the two. Although there were several significant correlations between the personal attitude factor (Factor 1) and implicit measures (WIT and IAT accuracy and RT performance bias), the size of the relations was modest ($rs < .17$).

WIT and FPST. We tested whether EF moderates the relationship between implicit bias and explicit bias with the structural
F6 equation model shown in Figure 6. This model uses the WIT and FPST latent performance bias variable and the Personal Attitudes factor from our factor analysis (Factor 1) as the explicit bias measure. The model without interactions fit the data well, $\chi^2(42) = 64.75$, $p = .014$, CFI = 0.964, RMSEA = 0.033. When taking this latent variable perspective, implicit bias significantly predicted explicit attitudes. However, there was no indication that this relation depended on EF, as shown by the nonsignificant Common EF $\times$ Implicit Bias interaction. The only other significant predictor was Updating-specific ability; relatively greater Updating-specific ability was associated with less negative explicit attitudes toward Blacks. It is difficult to determine what this

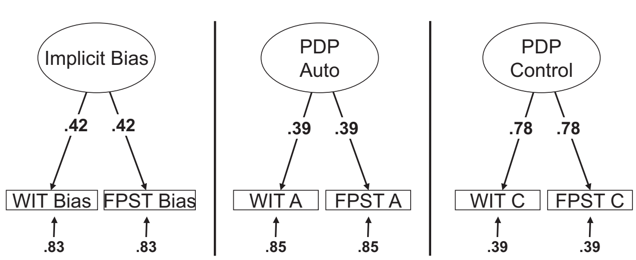

Figure 2. Three latent variable models for bias measures. Numbers on straight arrows are standardized loadings, and numbers at the end of arrows are residual variances. In each model, analogous measures (bias scores, PDP auto (A) scores, or PDP control (C) scores) from the Weapons Identification Task (WIT) and First Person Shooter Task (FPST) load on a single latent variable. Because each factor has only two indicators, path loadings (on standardized scores) are constrained to be equal to ensure statistical identification. Boldface type indicates $p < .05$.

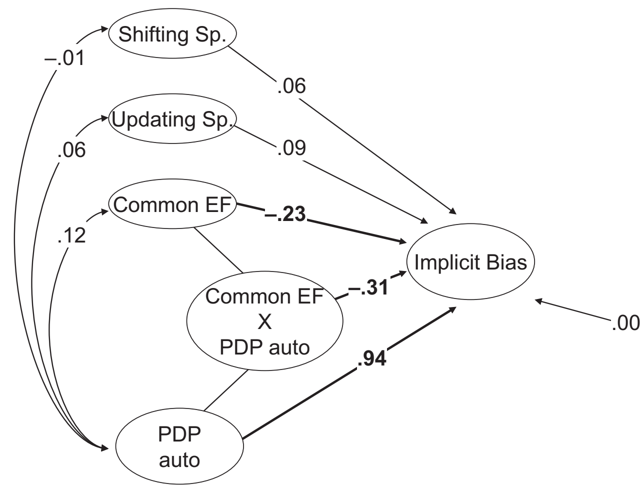

Figure 3. Structural equation model predicting implicit bias (latent variable constructed from accuracy bias scores on the WIT and the FPST) with Common EF, PDP auto (latent variable constructed from PDP auto scores on the WIT and FPST), and the interaction of PDP auto and Common EF. Ovals indicate latent variables (individual indicators not shown for simplicity). Numbers on straight arrows are standardized path coefficients, numbers on curved double-headed arrows are correlations, and numbers at the end of arrows are residual variances. The residual variance for implicit bias was constrained to be 0.0 to prevent a nonsignificant negative residual variance. Boldface type indicates $p < .05$.

relationship might signify as the Updating-specific factor is probably the least well-understood in terms of its primary sources of variance (Miyake & Friedman, 2012), with candidates being working memory gating by the basal ganglia, retrieval from long-term memory, and perhaps capacity limitations.[8] Fn8

IAT. Similar effects were observed when predicting explicit attitudes from IAT bias (see Figure 7). The model without inter- F7
actions fit the data well, $\chi^2(32) = 47.55$, $p = .038$, CFI = 0.975, RMSEA = 0.032. IAT bias significantly predicted explicit attitudes, but this relation was not moderated by any aspect of EF. As in the WIT/FPST model, greater Updating-specific ability was associated with less negative explicit attitudes (as should be expected given that these two variables are the same across the models depicted in Figures 6 and 7). However, in contrast to the WIT/FPST model, the path from Shifting-specific to Personal

[8] Though we report results using factor scores extracted from the exploratory factor analysis presented in Table A1, we obtained similar results when using scale scores for the measures with items included in the factor analysis. For example, when we ran the models shown in Figures 6, 7, 8, and 11 with the Attitudes Toward Blacks scale (ATB; Brigham, 1993) instead of the Personal Attitudes factor score, conclusions were the same (implicit bias significantly predicted explicit attitudes, but there was no indication that this relation depended on EF). The pattern in these models was also the same when using the funding allocation bias score on its own instead of as part of the Personal Attitudes factor score. Finally, the models pictured in Figures 8, 10, 11, and 12 with Internal and External Motivation to Control Prejudice Scale Scores (Plant & Devine, 1998) used in place of the factor scores for those constructs also resulted in similar patterns, though some paths dropped from significant to marginally significant. The overall similarity of results with the factor scores and scale scores suggests that the factor scores appear to be suitable composites of the constructs tapped by the individual explicit measures.

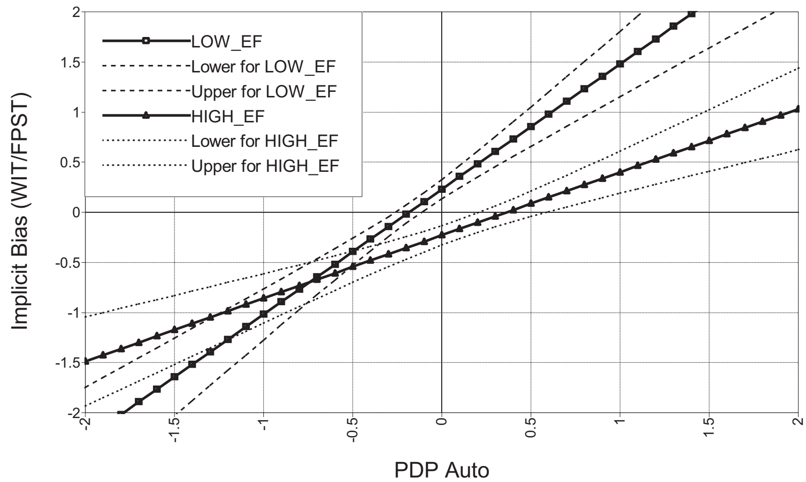

Figure 4. Plot of simple slopes for the PDP Auto × Common EF interaction predicting implicit bias shown in Figure 3, created with the plot function in Mplus. Solid lines for low and high Common EF represent one standard deviation below or above the mean, with dashed and dotted lines depicting the 95% confidence intervals.

Attitudes also reached significance, suggesting that better Shifting-specific ability was also associated with less negative explicit attitudes.

Motivation to Control Prejudice, EF, and Implicit and Explicit Bias (H5 and H6)

WIT and FPST. Our last set of analyses focused on relations among motivations to control prejudice, EF, and implicit bias. We have already established that EF weakens the relation between automatic processes and implicit bias (H2, Figure 3), showing that domain-general control abilities affect the expression of implicit racial bias. Our interest here is in the role of race-specific control concerns, in the form of motivations to control prejudice. We had two specific predictions: that motivation to control prejudice contributes to variability in bias above and beyond the contribution of EF (H5), and that EF and motivation to control prejudice interact to affect bias (H6). We evaluated these hypotheses for both implicit and explicit bias. If both the domain-specific concern about appearing biased in this particular context and the general ability to implement control are needed to modulate the impact of biased automatic associations, implicit and explicit bias should be lowest among those who have both high motivation to control prejudice and high levels of EF. This may be especially true for internal sources of motivation (cf. Amodio et al., 2008, 2003; Devine et al., 2002; Gonsalkorale et al., 2011).

F8 These hypotheses are tested in the model depicted in Figure 8 using the Personal Attitudes factor from our factor analysis (Factor 1) to represent explicit bias, and the Internal and External Motivation factors (Factors 2 and 3, respectively) for the two motivation variables. The model without interactions fit the data well, $\chi^2(56) = 84.70$, $p = .008$, CFI = 0.966, RMSEA = 0.033. Consistent with H5, Internal Motivation independently predicted both implicit bias and Personal Attitudes. In both cases, higher levels of internal motivation were associated with lower levels of racial bias. Consistent with past research (Devine et al., 2002), external motivation was positively associated with negative personal attitudes, indicating that higher levels of external motivation to control prejudice are associated with higher levels of explicit bias (controlling for internal motivation). External motivation did not directly predict implicit bias, but did interact with Common EF. This interaction is depicted in Panel A of Figure 9 and F9
provides some support for H6. For individuals with lower Common EF, higher external motivation is associated with higher levels of implicit bias, (simple slope $b = 0.43$, $SE = 0.20$, $z = 2.17$, $p = .030$), but this relation weakens at higher levels of Common EF, (simple slope $b = -0.25$, $SE = 0.18$, $z = -1.39$, $p = .165$). This relation suggests that the tendency for external motivation to be associated with greater bias is particularly true when Common EF is low. Note that Common EF moderated the effects of external motivation only for implicit but not explicit bias.

The only other significant effects in this model were negative relations of Shifting-specific and Updating-specific abilities with negative personal attitudes. These indicate that higher levels of Shifting-specific and Updating-specific abilities are associated with lower levels of explicit bias, as was the case in Figure 7. The direct relation between Common EF and implicit bias was not significant in this model, but its size and direction were the same as the model in Figure 3 (in this latter model, the PDP auto factor explained a good deal of variance in performance bias, which resulted in less error variance and tighter confidence intervals). Unlike the Shifting-specific and Updating-specific factors, Common EF did not predict explicit bias.

Together, these results confirm H5 in showing the effect of domain-specific internal motivation to control prejudice responses on racial bias. Consistent with theorizing that motivation to control prejudice should reduce its expression, higher internal motivation predicted both lower levels of implicit and explicit bias independent from EF ability. Support for this hypothesis was more mixed for external motivation. There was no direct effect of external

motivation to control prejudice on implicit bias but it did predict more negative explicit racial attitudes. Moreover, instead of exerting a direct effect on implicit bias, the effect of external motivation was moderated by Common EF, with the tendency toward greater bias among those with higher external motivation observed in past research (Amodio et al., 2008, 2003; Devine et al., 2002; Gonsalkorale et al., 2011) weakening as Common EF increased. By contrast, the benefits of high internal motivation to control prejudice did not depend on Common EF. In this way, the results show an intriguing dissociation between internal and external sources of motivation to control prejudice, with internal motivation having a beneficial effect on implicit bias regardless of domain-general EF abilities. Because they are personally endorsed, internal sources of motivation are arguably more well practiced and thus may require fewer cognitive resources to implement (Amodio et al., 2006, 2008; Devine et al., 2002; Monteith, 1993; Monteith, Lybarger, & Woodcock, 2009), helping to uncouple them from overall EF abilities. By contrast, responding to social perceptions of appropriate behavior may be more resource dependent, causing the implementation of external motivations to rely more strongly on general EF abilities. Our results suggest this is more true for
Fn9 implicit than explicit bias.[9]

Under most extant theorizing, attempts to control implicit bias are thought to operate primarily through the recruitment of effortful processing rather than through changes in automatic processes.

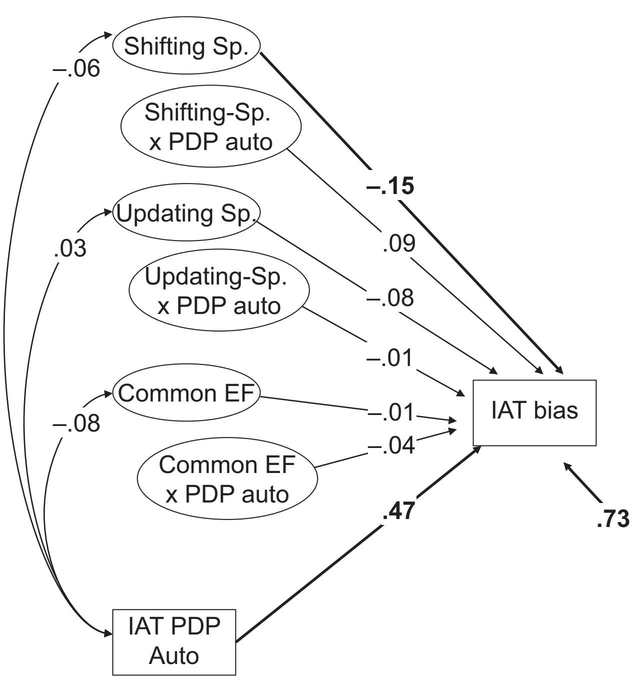

Figure 5. Structural equation model predicting implicit bias in the IAT with executive functions (EF), PDP auto scores, and their interactions. Ovals indicate latent variables (individual indicators not shown for simplicity), rectangles indicate observed measures. Numbers on straight arrows are standardized loadings, numbers on curved double-headed arrows are correlations, and numbers at the end of arrows are residual variances. This model is analogous to the one shown in Figure 3, but because IAT shows different relations with the EF components, we depict all three possible interaction effects. Boldface type indicates $p < .05$.

Table 5
Standardized Path Coefficients From EF Latent Variables to PDP Control and Auto

Bias variables	Common EF	Updating—specific	Shifting—specific
WIT PDP control	.41*	−.06	−.05
FPST PDP control	.45*	−.04	.11
IAT PDP control	.08	.20*	−.14*
WIT PDP auto	−.02	.07	.08
FPST PDP auto	.06	.00	−.10
IAT PDP auto	−.09	.06	−.06

Note. Each row represents a separate structural equation model in which the bias variable was regressed on the three EF latent variables.
* $p < .05$.

We would thus expect the effects of EF and motivation observed in Figure 8 to occur primarily through effects on PDP control but not auto. We tested this idea using the model depicted by Figure 10. Here Internal Motivation, External Motivation, the three aspects of EF, and the Common EF × Motivation interactions were used to predict both PDP control and PDP auto. The model without interactions fit the data well, $\chi^2(67) = 93.93$, $p = .017$, CFI = 0.969, RMSEA = 0.029. As expected, there were no effects of any aspect of EF or motivation on PDP auto. By contrast, both Common EF and Internal Motivation were independent predictors of PDP control. External Motivation, however, had no unique effect on the implementation of control. Note also the absence of EF × Motivation interactions on control, suggesting that when overall performance on implicit bias tasks is decomposed into estimates of underlying processes, the influence of motivations to control prejudice are not moderated by EF ability. F10

IAT. The same model predicting implicit and explicit bias from EF and motivation to control prejudice is shown in Figure 11, F11 using the IAT *D*-score as the measure of implicit bias. The model without interactions fit the data well, $\chi^2(44) = 59.48$, $p = .060$, CFI = 0.981, RMSEA = 0.027. Consistent with H5, and replicating the pattern seen with the WIT and FPST, Internal Motivation was an independent predictor of IAT bias, with less bias expressed by individuals higher in Internal Motivation. However, the IAT analyses also showed an independent effect of External Motivation: Higher levels of external motivation were associated with more implicit bias. As with the WIT/FPST analysis and consistent with H6, External Motivation also interacted with Common EF. This interaction, depicted in Panel B of Figure 9, was similar to that obtained with WIT and FPST: The positive relation between external motivation to control prejudice and IAT bias

[9] Past research suggests that internal and external motivation interact to affect the expression of prejudice (i.e., evaluative associations), with the lowest level of bias among those high in internal motivation but low in external motivation (Amodio et al., 2006, 2008; Devine et al., 2002; Gonsalkorale et al., 2011; Plant & Devine, 1998). We might not expect such an interaction in the current data, given that the implicit measures we used were structured to reflect stereotyping, and as argued by Amodio, Devine, and Harmon-Jones (2008), the interaction of these two facets of motivation should predict implicitly measured evaluations rather than stereotypes. Indeed, when models were modified to include the interaction between internal and external motivation, this interaction failed to predict WIT/FPST or IAT bias.

APA NLM
tapraid5/z2g-perpsy/z2g-perpsy/z2g00215/z2g4288d15z | xppws | S=1 | 1/7/15 | 3:54 | Art: A-2013-0804

Table 6
Intercorrelations Among Factor Scores of Explicit Measures and Individual Implicit Measures

	Extracted factor			
Measure	1. Personal attitude	2. Internal motivation	3. External motivation	4. Cultural stereotypes
WIT accuracy performance bias	.16*	−.01	.05	.06
FPST accuracy performance bias	.05	−.08	−.05	.04
IAT accuracy performance bias	.16*	−.04	.04	.10
WIT RT performance bias	.16*	−.03	−.02	.02
FPST RT performance bias	.07	.00	.01	−.03
IAT D-score performance bias	.16*	−.03	.06	.06
WIT PDP auto	.09	.02	.06	.08
FPST PDP auto	.04	−.07	−.05	.05
IAT PDP auto	.11*	.02	.06	.07
WIT PDP control	−.18*	.12*	.05	−.01
FPST PDP control	−.09	.02	.02	.03
IAT PDP control	−.17*	.13*	.03	−.02

Note. Higher values indicate greater racial bias in accuracy, reaction time, and PDP auto, higher mean control, more negative Personal Attitudes, higher internal and external motivation to control prejudice, and stronger perceived stereotypes. Correlations are maximum likelihood estimates (from Mplus) based on all data, adjusted for missing observations. Total $N = 481$. RT = reaction time.
* $p < .05$.

seen at low Common EF (simple slope $b = 0.31$, $SE = 0.11$, $z = 2.92$, $p = .004$) was eliminated in those with higher Common EF (simple slope $b = -0.01$, $SE = 0.10$, $z = -0.09$, $p = .931$).

As with the WIT/FPST analyses, we computed a model in which motivation and EF predicted IAT PDP control and auto to assess whether motivation and EF primarily influence control but not auto (see Figure 12). The model without interactions fit the data well, $\chi^2(44) = 63.47$, $p = .029$, CFI = 0.969, RMSEA = 0.030. Replicating the pattern seen with WIT/FPST, Internal Motivation

F12

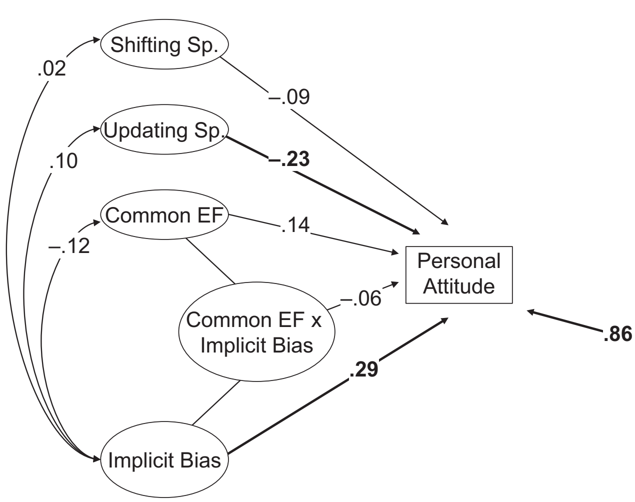

Figure 6. Structural equation model predicting Personal Attitudes (Factor 1 in Table A1) with executive functions (EF), implicit bias, and the interaction of bias with Common EF. Ovals indicate latent variables (individual indicators not shown for simplicity), rectangles indicate observed measures. Numbers on straight arrows are standardized loadings, numbers on curved double-headed arrows are correlations, and numbers at the end of arrows are residual variances. Boldface type indicates $p < .05$.

was associated with greater control. The only other significant predictor was Shifting-specific ability. As noted previously (see Table 5), greater Shifting-specific ability is associated with *less* control in the IAT.

Discussion

The present study sought to provide a comprehensive assessment of the relationship between EF and implicit racial bias. We begin our discussion of the results with an examination of the primary methodological and theoretical contributions made by this study and how they relate to our primary hypotheses concerning the ways in which EF moderates expression of implicit bias (H1 and H2) and what this implies about the nature of control over implicit bias. We then discuss effects related to the ancillary predictions and other implications of our results. Because of the large number of hypotheses assessed, we also summarize results with respect to each specific hypothesis in Table 7. T7

The basic question of whether cognitive control is involved in performance on laboratory measures of implicit racial bias has been the focus of researchers' attention for some time now (see Amodio & Mendoza, 2010; Bartholow et al., 2006, 2012; Conrey et al., 2005; Klauer & Mierke, 2005; Klauer et al., 2007; Payne, 2001, 2005; Schlauch et al., 2009; Siegel et al., 2012). Relative to these previous investigations, the approach taken in the current work provided a number of significant methodological advances. In particular, the current study is arguably the most comprehensive assessment of both EF and implicit racial bias yet attempted. Informed by recent advances in understanding the latent structure of EF (e.g., Miyake & Friedman, 2012), the study involved assessment of three distinct yet related facets of EF, each indicated by performance on multiple, well-validated laboratory tasks. This allowed the derivation of latent factors representing the "unity and diversity" of EF abilities (see Miyake et al., 2000; Teuber, 1972). Likewise, participants in the current study completed three different laboratory measures of implicit racial bias, each designed to

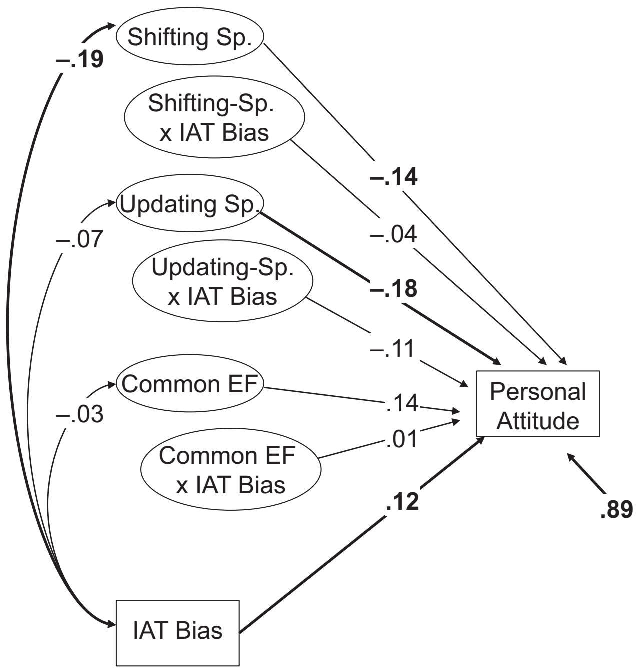

Figure 7. Structural equation model predicting Personal Attitudes (Factor 1 in Table A1) with executive functions (EF), implicit bias in the IAT, and their interactions. Ovals indicate latent variables (individual indicators not shown for simplicity), rectangles indicate observed measures. Numbers on straight arrows are standardized loadings, numbers on curved double-headed arrows are correlations, and numbers at the end of arrows are residual variances. This model is analogous to the one shown in Figure 6, but because IAT shows different relations with the EF components, we depict all three possible interaction effects. Boldface type indicates $p < .05$.

assess the strength of automatic associations between young Black men and danger but in somewhat different ways. Thus, unlike most previous attempts to link EF and performance on measures of implicit racial bias, in which both EF and implicit racial bias have been represented by performance on only one or a small handful of tasks (Klauer & Mierke, 2005; Mierke & Klauer, 2003; Payne, 2005; Siegel et al., 2012; but see Klauer et al., 2010), the design of the current study allowed for tests of whether individual differences in theoretically distinct facets of EF are associated with expression of racial bias at the level of latent variables. This provides information on the nature of control involved in the performance of implicit racial bias tasks, as well as the degree to which this control is consistent across multiple measures of bias. Furthermore, the relatively large number of participants in the current study, drawn from multiple geographic locations, increases the representativeness of the sample.

The primary finding from this study, pertinent to H1, is that despite a fair degree of overlap in estimates of general controlled processing during the implicit bias tasks (i.e., PDP control; *r*s ≥ .29; see Table 4), the impact of the three components of EF varied across the tasks. Our results indicate that bias expressed on the WIT and FPST is affected by individual differences in Common EF. Miyake and Friedman (2012) argue that Common EF represents the ability to actively maintain and use task goals to influence lower-level processing, particularly in the face of interference. Notably, this factor involves more than inhibition; the latent variable has significant loadings from all nine of the EF tasks used in the current study. Thus, whereas much past research has focused on the role of inhibition in modulating bias expression (Devine, 1989; Gilbert & Hixon, 1991; Govorun & Payne, 2006; Macrae et al., 1994; Monteith et al., 1998; Payne, 2005; Richeson & Shelton, 2003; von Hippel, 2007), our results point to the possibility that cognitive control of implicit bias may involve more than simply the ability to inhibit prepotent responses by virtue of the contribution of performance on the shifting and updating tasks to the Common EF variable. This more expansive view of the mechanisms involved in control of implicit racial bias converges with that argued by others (e.g., Conrey et al., 2005; Payne, 2001).

In contrast to the findings for WIT and FPST, Common EF played a much less prominent role in the control of racial bias in the IAT. Rather, we observed correlations of IAT *D*-score and PDP control with Shifting-specific ability. This dissociation in the types of EF related to the different bias tasks speaks to important differences in the cognitive processes involved in task performance for the various measures. To date, this issue has received relatively little consideration (but see Klauer et al., 2010). Although the three bias tasks were designed to measure the same associations (i.e., stereotypes linking young Black men with danger) in roughly the same way (i.e., speeded judgments), they differ in several key ways. For example, the WIT and IAT separate the racial category information (conveyed via faces) from the stereotype-relevant content (conveyed via stereotypic images or words) whereas the two are contained in a single stimulus in the FPST (people of different races holding objects that differ in stereotypicality). The WIT is a sequential priming task, but FPST is not, and the IAT alternates stimuli and tasks on a trial-by-trial

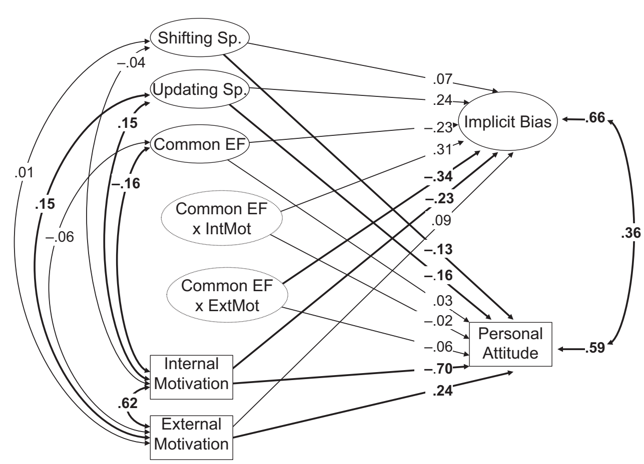

Figure 8. Structural equation model predicting implicit bias (latent variable constructed from bias scores on the Weapons Identification Task and the First Person Shooter Task) and Personal Attitudes (Factor 1 in Table A1) with executive functions (EF), internal and external motivation to control prejudice (Factors 2 and 3 in Table A1), and the interactions of these motivation scores with Common EF. Ovals indicate latent variables (individual indicators not shown for simplicity), rectangles indicate observed measures. Numbers on straight arrows are standardized loadings, numbers on curved double-headed arrows are correlations, and numbers at the end of arrows are residual variances. Boldface type indicates $p < .05$.

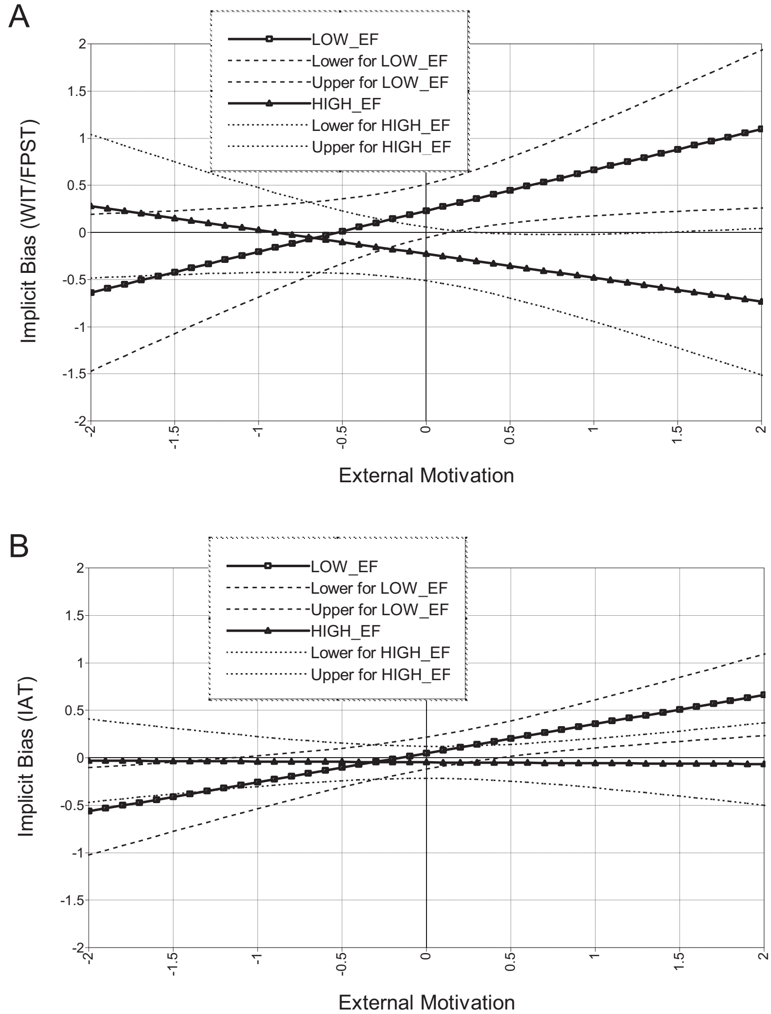

Figure 9. Plot of simple slopes for the External Motivation × Common EF interactions shown in Figures 8 and 11, using the plot function in Mplus. Solid lines for low and high Common EF represent one standard deviation below or above the mean, with dashed and dotted lines depicting the 95% confidence intervals. Panel A shows interaction of external motivation and Common EF predicting the latent implicit bias score from Weapons Identification Task and First Person Shoot Task. Panel B shows the same interaction predicting IAT *D*-scores; although the 95% confidence intervals overlap in this latter figure, the interaction was significant according to the chi-square difference test ($p < .05$).

(and block-by-block) basis. The WIT and FPST involve tight response deadlines and performance is accuracy-based, whereas the IAT does not use a response deadline and performance is RT-based. Given these structural differences, it should not be surprising that they draw on different aspects of EF, but it is not clear how fully researchers consider the influence of these features when selecting a task to use.

Interestingly, our analyses show that not only do the bias tasks differ in terms of the aspects of EF to which they relate, but they also differ in terms of how those EF facets relate to PDP estimates of control drawn from the tasks. Specifically, Common EF was negatively associated with WIT/FPST bias (see Figure 3) and positively associated with PDP control drawn from those tasks (see Figure 10). In other words, better Common EF ability predicted *less* WIT/FPST bias and *more* general control during those tasks. This could suggest that PDP control reflects the influence of cognitive processes represented by the Common EF variable. In contrast, Shifting-specific ability was negatively associated with both IAT bias (see Figure 5) and PDP control in the IAT (see Figure 12). In other words, better Shifting-specific ability predicted both *less* IAT bias and *less* general control during the task. At first blush this pattern seems contradictory, but the resolution may lie in the types of control being captured in the Shifting-specific EF and PDP control variables in the IAT. Klauer and colleagues (e.g., Klauer et al., 2010; Mierke & Klauer, 2003) have argued that the link between Shifting-specific ability and IAT performance reflects method variance that is independent of the measure's content. Specifically, they suggest that the IAT's relation to Shifting-specific ability reflects the inherent need to switch task sets from trial-to-trial between a concept categorization task

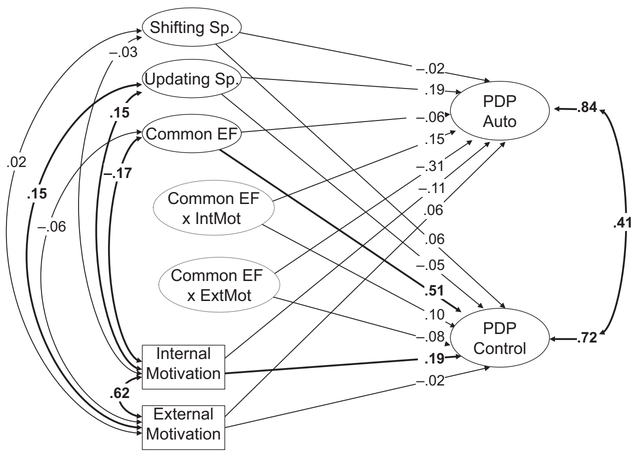

Figure 10. Structural equation model predicting PDP auto and PDP control (latent variables constructed from the PDP scores on the Weapons Identification Task and the First Person Shooter Task; residuals for each test allowed to correlate) with executive functions (EF), internal and external motivation to control prejudice (Factors 2 and 3 in Table A1), and the interactions of these motivation scores with Common EF. Ovals indicate latent variables (individual indicators not shown for simplicity), rectangles indicate observed measures. Numbers on straight arrows are standardized loadings, numbers on curved double-headed arrows are correlations, and numbers at the end of arrows are residual variances. Boldface type indicates $p < .05$.

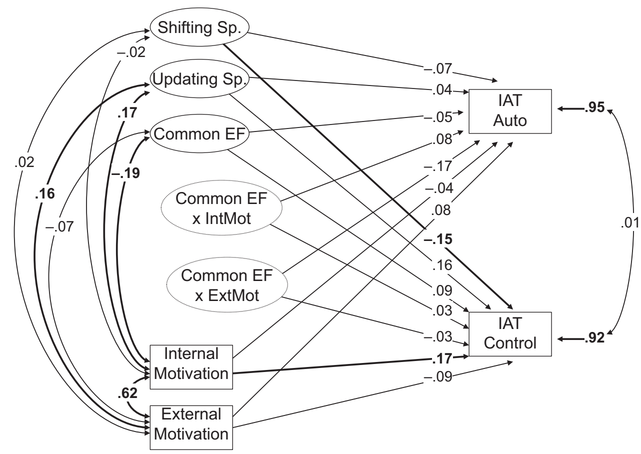

Figure 12. Structural equation model predicting PDP auto and PDP control for the IAT with executive functions (EF), internal and external motivation to control prejudice (Factors 2 and 3 in Table A1), and the interactions of these motivation scores with Common EF. This model is analogous to the one shown in Figure 10. Ovals indicate latent variables (individual indicators not shown for simplicity), rectangles indicate observed measures. Numbers on straight arrows are standardized loadings, numbers on curved double-headed arrows are correlations, and numbers at the end of arrows are residual variances. Boldface type indicates $p < .05$.

(e.g., race) and an attribute categorization task (e.g., word valence). Shifting-specific ability should thus be implicated in the IAT regardless of the type of associations being measured. Consistent with this interpretation, Klauer and colleagues observe relations between Shifting-specific ability and IAT performance across IATs assessing different types of associations (e.g., the

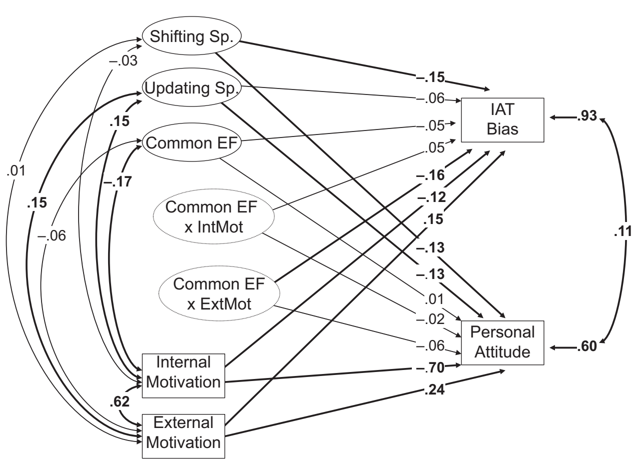

Figure 11. Structural equation model predicting implicit bias on the IAT and Personal Attitudes (Factor 1 in Table A1) with executive functions (EF), internal and external motivation to control prejudice (Factors 2 and 3 in Table A1), and the interactions of these motivation scores with Common EF. This model is analogous to the one shown in Figure 8. Ovals indicate latent variables (individual indicators not shown for simplicity), rectangles indicate observed measures. Numbers on straight arrows are standardized loadings, numbers on curved double-headed arrows are correlations, and numbers at the end of arrows are residual variances. Boldface type indicates $p < .05$.

relative positivity of flowers and insects, political parties, and racial groups). We similarly show that higher Shifting-specific ability is associated with reduced performance bias in the IAT. By contrast, estimates of PDP control reflect the ability to behave as intended (Payne, 2001), which presumably requires both switching from race classification to valence classification and simultaneously avoiding the influence of racially stereotypic associations. What is interesting in our data is that Shifting-specific ability and IAT PDP control were negatively related. We speculate that this may indicate a trade-off in the type of control being implemented, with those implementing more general task switching ability deploying less racial association-specific PDP control. Given the popularity of the IAT, further exploration of the impact of different types of control on task performance would be informative (see Hilgard, Bartholow, Dickter, & Blanton, 2014).

Beyond these simple, direct associations, our results also speak to the issue of EF moderating the influence of automatic associations on bias expression (H2). As illustrated in Figure 4, for individuals relatively low in EF, bias in PDP auto translates more directly into performance bias on the implicit measures; individuals relatively high in EF, however, seem better able to reduce the influence of automatic processes on their behavior. By successfully limiting the biasing impact of these automatic associations, individuals higher in EF are less affected by the incidental racial features of the implicit measures that are objectively irrelevant to task goals (but critical for measuring racial bias). There was also little difference in bias as a function of EF when PDP auto was low, suggesting that level of EF will have less impact on the expression of bias in the absence of a stereotypic impulse. By contrast, when stereotypic response tendencies are higher (as suggested by higher PDP auto), EF becomes more critical, such that those higher in EF are better able to avoid behaving in line with their automatic tendencies.

Table 7
Summary of Results

Number	Hypothesis	WIT/FPST result	IAT result
H1	Performance on implicit bias tasks reflects both EF and automatic processes.	SUPPORTED: Greater bias in PDP auto predicts higher performance bias. Greater Common EF predicts lower performance bias (Figure 3).	SUPPORTED: Greater bias in PDP auto predicts higher performance bias. Greater Shifting-specific ability predicts lower performance bias (Figure 5).
H2	EF moderates the effect of automatic processes on implicit bias, with automatic processes more directly driving performance bias among those with weaker EF abilities.	SUPPORTED: Common EF and PDP auto interact, with the positive relation between PDP auto and bias weakened at higher levels of EF (Figures 3 and 4).	NOT SUPPORTED: No interaction between any aspect of EF and PDP auto (Figure 5).
H3	EF should relate to control but not automatic processes.	SUPPORTED: Greater Common EF ability was associated with greater control. As predicted, no relations between EF and PDP auto (Table 5).	MIXED: Greater Updating-specific ability was associated with greater control, but greater Shifting-specific ability was associated with *less* control. As predicted, no relations between EF and PDP auto (Table 5).
H4	EF moderates the relation between implicit and explicit bias, with a stronger implicit-explicit relation when EF is low.	NOT SUPPORTED: Negative Personal Attitudes (Factor 1) were positively predicted by stronger implicit bias and negatively by Updating-specific ability, but Common EF did not interact with implicit bias (Figure 6).	NOT SUPPORTED: Negative Personal Attitudes (Factor 1) were positively predicted by stronger implicit bias and negatively by Updating-specific ability, but no aspects of EF interacted with implicit bias (Figure 7).
H5	Bias will be negatively associated with internal motivation but positively associated with external motivation, even when controlling for EF.	SUPPORTED for internal motivation; MIXED for external motivation: Greater internal motivation was associated with less negative implicit and explicit bias. External motivation was unrelated to implicit bias, and was associated with more negative Personal Attitudes (Figure 8).	SUPPORTED for internal and external motivation: Greater internal motivation was associated with less negative implicit and explicit bias. Greater external motivation was associated with more negative implicit and explicit bias (Figure 10).
H6	EF moderates the effect of motivation to control prejudice on implicit bias, with motivation more effectively translating into decreased bias when EF is high.	NOT SUPPORTED for internal motivation; SUPPORTED for external motivation: No Common EF and internal motivation interaction. Common EF did interact with external motivation in predicting implicit bias. The positive relation between external motivation and implicit bias is strongest for people who have low Common EF (Figures 8 and 9).	NOT SUPPORTED for internal motivation; SUPPORTED for external motivation: No Common EF and internal motivation interaction. Common EF did interact with external motivation in predicting implicit bias. The positive relation between external motivation and implicit bias is strongest for people who have low Common EF (Figures 9 and 10).

Here, too, there was divergence in terms of the roles played by specific EF abilities in determining bias on the WIT/FPST versus the IAT. For the WIT/FPST latent bias variable, Common EF both predicted bias directly and moderated the effect of automatic associations on bias, suggesting that Common EF attenuates the influence of automatic associations on performance (see also Payne, 2005). By contrast, the effect of EF on IAT performance bias was limited to a direct effect of Shifting-specific ability on IAT *D*-score. Furthermore, in our data, we see no evidence that the impact of Shifting-specific ability on IAT performance interacts with automatic race-based associations (i.e., the PDP auto bias score). As a whole, this pattern suggests that, although Shifting-specific ability affects performance on the IAT, it is not related to participants' ability to intentionally control the influence of racially biased associations.

Relations Between Measures of Implicit Racial Bias

The methodological innovations in our design permit a number of additional theoretical advances to be considered. First, although not a major focus of the current study, our analyses revealed that the three measures of racial bias used here showed only weak associations with one another. Although this perhaps could be expected at the level of overall performance bias (see Fazio & Olson, 2003), which arguably is contaminated by differences in structural features and the types of controlled processes invoked across the different tasks, similar patterns emerged for estimates of the extent to which automatic processes contribute to performance (i.e., PDP auto). Even the strongest correlation in PDP auto estimates, between WIT and FPST, indicates that these two measures share only 4% of their variance. Although this overlap was strong enough to allow us to extract a latent variable and relate it to our EF model, more work must be done to clarify the precise mental content assessed by these implicit measures, and to determine the extent to which differences in what the tasks measure contributes to their limited association. Furthermore, this lack of correspondence across measures suggests that researchers must use caution when making decisions concerning which laboratory measure(s) of racial bias to use.

APA NLM
tapraid5/z2g-perpsy/z2g-perpsy/z2g00215/z2g4288d15z | xppws | S=1 | 1/7/15 | 3:54 | Art: A-2013-0804

Relations Between Implicit and Explicit Racial Bias

In addition to addressing critical theoretical issues about the contribution of control to performance on implicit bias tasks, as well as the nature of that control, our data speak to the issue of how implicit and explicit measures of bias are related. Similar to the findings of Cunningham, Preacher, and Banaji, (2001) (see also Hofmann, Gawronksi, Gschwender, Le, & Schmitt, 2005), we observed a relationship between the two types of measures that was more plainly evident when implicit bias was estimated using a latent variable rather than measure-specific indices of performance bias. However, we found no support for the prediction that participants with lower EF ability show tighter correspondence between implicit and explicit measures of racial bias due to poorer control over performance on implicit measures (H4). This prediction can be thought of as an individual differences extension of dual process model assumptions that explicit measures are more sensitive to controlled processes. To the degree that the control referred to in dual process models depends on EF, we reasoned that people with greater EF would have more capacity to control explicit responses than those with lower EF, resulting in weaker relations between implicit and explicit measures for those with higher EF. Instead, the relationship between implicit and explicit measures did not depend on EF. Thus, unlike past studies that have changed the implicit-explicit relation through situational manipulations of cognitive control (Payne et al., 2008; Ranganath et al., 2008), individual differences in EF did not show a comparable moderating effect.

Motivation to Control Prejudice

The literature on racial attitudes has emphasized the influence of specific motivations to control prejudice (Amodio et al., 2008; Dunton & Fazio, 1997; Glaser & Knowles, 2008; Maddux et al., 2005; Park, Glaser, & Knowles, 2008; Plant & Devine, 1998). To date, however, there has been little consideration of how general cognitive ability and domain-specific motivations operate together (but see Amodio et al., 2008; Payne, 2005). Our design and large sample permitted us to assess their simultaneous effects to determine whether they independently influence implicit and explicit expression of bias (H5). Our results showed that high internal motivation predicts less bias in WIT/FPST, IAT, and Personal Attitudes, as well as greater control on implicit bias tasks (as assessed with PDP control). This beneficial effect of internal motivation does not depend on EF (i.e., internal motivation does not interact with Common EF). By contrast, high external motivation is associated with greater bias in the IAT and more negative personal attitudes, a pattern that weakens (for the implicit measure) with increases in Common EF.

There are several theoretical implications of these results. They first suggest that cognitive ability (i.e., EF) and motivation are distinct, independent predictors of bias expression, at least when motivation arises from internal factors. People who are internally motivated to avoid prejudiced responding are theorized to have internalized egalitarian standards and apply them consistently across situations (e.g., Devine et al., 2002; Monteith et al., 2009; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999). Our results suggest the application of this motivational drive is relatively unaffected by variation in EF. By contrast, the impact of external motivation appears to be more resource dependent. It has been suggested that, whereas internal motivation produces negative self-directed affect if behavior falls short of internalized egalitarian standards, feeling externally pressured to inhibit bias can produce resentment and frustration, which appears to not successfully translate into reduced bias. In our data, participants high in external motivation show evidence of racial bias on both implicit and explicit measures. Those individuals high in EF, however, are able to reduce the expression of bias when completing the implicit tasks, presumably due to their increased ability to adhere to overarching task goals (e.g., to correctly identify guns and tools, regardless of the preceding racial prime).

In many ways this pattern makes intuitive sense. External motivation reflects a concern with appearing prejudiced (perhaps even awareness of one's own socially inappropriate attitudes), whereas internal motivation reflects differences in the desire to be—not simply to appear—egalitarian. The differences in these constructs may account for the divergence in relations between EF and the two forms of motivation: whereas belief in the value of egalitarianism might in itself be enough to reduce bias expression, a desire to *appear* nonprejudiced for external reasons arguably requires more effort when the situation calls for bias control—the very kind of situation in which EF should be important for modifying behavior. As noted by previous scholars, sources of internal motivation are arguably more well-practiced and thus may require fewer cognitive resources to implement (Monteith et al., 2009; Moskowitz et al., 1999), potentially uncoupling their effects from EF abilities per se and permitting better implementation of task-specific control. This pattern was evident here by the fact that internal motivation was directly related to PDP control (see also Amodio et al., 2008) but external motivation was not.

Motivation and EF were unrelated to PDP auto estimates, in line with the assumption that PDP auto reflects processes immune from the influence of intentions (Payne, 2001). Others, however, have shown that internal motivation is associated with less activation of automatic associations (Gonsalkorale et al., 2011), a finding suggesting that those high in internal motivation either lack biased automatic associations or possess the ability to automatically inhibit them (see Amodio et al., 2003; Devine et al., 2002). Intuitively, the findings and interpretation in these studies seem to fit with our results (discussed above) showing that the influence of internal motivation is not moderated by EF (suggesting that it is not resource dependent). It is therefore not clear to us why the results from Gonsalkorale et al. (2011) differ from those observed here with respect to the relationship between motivation and PDP auto. One possibility is the use of different measures of automatic processes in the two studies (PDP vs. Quad Model). Given the theoretical and practical implications of knowing whether variability in racially biased behavior is primarily affected by differences in control, or whether differences in automatic processes also contribute, this issue merits further investigation.

The Meaning of PDP Auto and Control Estimates

As highlighted in the Discussion section thus far, one remarkable aspect of our findings is the relative lack of correspondence between the PDP estimates derived from our three implicit tasks. Despite our efforts to equate the stereotype content assessed by the measures, there was relatively little overlap in estimates of PDP auto, and while there were significant correlations in PDP control, these estimates actually related to very different forms of EF when examined in the context of our latent factor models. Together, these findings reinforce many of the basic lessons gleaned from past research using the PDP approach. That is, although the PDP formulas offer a powerful technique for examining the relative contribution of controlled and automatic processes within a given task, those estimates are constrained by the particular content and structure of the task. The only way to obtain domain general estimates of control is to use completely separate tasks that are known to measure these targeted processes (as was done here). One cannot simply assume that the PDP estimates observed in different tasks reflect similar underlying processes.

This conclusion also highlights a striking limitation in the field. There are currently no agreed upon process-pure measures of automatic racial associations, so while it was possible for us to measure control abilities independent from the bias tasks, we could not do the same for automatic processes. Because of this limitation, it is hard to interpret the meaning of the stronger relationship we observed between PDP auto and performance bias in the WIT/FPST versus the IAT. These differences could reflect a differential contribution of automatic processes to responding in the different tasks, or simply greater measurement overlap for WIT and FPST. Both PDP auto and performance bias are computed using error rates for the WIT/FPST, but only the PDP estimate is based on this accuracy data for the IAT (for which performance bias is scored primarily with RTs). Future research should examine whether the tasks we used differ in the degree to which they are influenced by automatic processes. Given the central importance of this issue to the field, identifying ways to quantify these automatic processes in a purer manner (as has been accomplished in the realm of EF) seems critical.

Another important implication of our work concerns the sufficiency of PDP control estimates to capture the control being implemented in frequently used implicit bias tasks. Our results indicate that it depends on the type of implicit bias task in which control is measured. For WIT and FPST, Common EF and PDP control were each associated with less bias, and Common EF and PDP control were themselves strongly related, suggesting that PDP control is tapping the influence of Common EF in these tasks.

To more explicitly evaluate this conclusion, we conducted an ancillary analysis in which both Common EF, PDP control, and the interactions of each with PDP auto simultaneously predicted the latent WIT/FPST performance bias measure. This was, in essence, the same model depicted in Figure 3, but with PDP control and the PDP Control × PDP Auto interaction added as predictors and allowed to correlate with the EF variables. Note that the estimates of PDP control and performance bias have a much higher degree of temporal and task specificity than do the EF measures and performance bias. Even if Common EF and PDP control reflect similar constructs, we therefore expected PDP control to emerge as the primary predictor of performance bias. Indeed this was the case; PDP control and the PDP Control × PDP Auto interaction predicted implicit bias, but Common EF—neither directly nor via an interaction with PDP auto—did not. This suggests that Common EF does indeed play a vital role in the type of control enacted during the WIT and FPST. In essence, the PDP control estimate derived from these tasks is a direct proxy for Common EF.

The relation of EF to PDP control is different for the IAT. As noted previously, in the IAT, Shifting-specific ability and PDP control were negatively correlated, but each was associated with less IAT bias, a pattern suggesting that EF and PDP control make independent contributions to IAT performance. As with WIT/FPST, we conducted an ancillary analysis including both EF and PDP control as predictors of IAT bias (i.e., a model similar to that in Figure 5 but adding the PDP control and PDP Control × PDP Auto interaction). Shifting-specific ability remained an independent predictor of IAT bias, but with Shifting-specific ability (as well as all other facets of EF) in the model, PDP control did not predict IAT bias. Consistent with earlier conclusions, there appears to be unique, nonoverlapping aspects of control contributing to IAT performance and fully capturing both the type and amount of control in the IAT might require independent measures of EF. To the question of whether PDP control is sufficient to represent the influence of Common EF in these tasks, if the goal is to determine the *type* of control recruited in these tasks, independent quantification of different aspects of EF is needed, but if the goal is to quantify the *amount* of control recruited, relying on PDP control may be sufficient.

Limitations and Suggestions for Future Research

Although the current study represents perhaps the most comprehensive examination to date of the relationship between EF and implicit racial bias, it was not without limitations. First, despite its relatively large size, the sample used in the current study likely represented a restricted range of responses in terms of both EF and racial bias. All participants were drawn from undergraduate populations at three relatively large universities, who arguably have both better cognitive abilities and less biased racial attitudes than the general population. This issue could help account for the fact that although our analyses revealed predicted patterns of association between EF and implicit bias, the magnitude of those associations generally was modest. Replication of the current findings in a sample drawn from a broader population would be extremely valuable.

A second limitation concerns differences in the structural features of the bias tasks used here, which likely contributed to the very modest associations among their bias estimates. While our use of tasks with different structural properties had the benefit of minimizing the influence of particular measurement characteristics, it is not surprising that tasks measuring bias in different ways produce estimates of bias that are not highly correlated with each other. However, to the extent that these tasks truly measure the same implicit biases, they ought to correlate for more than just methodological reasons. Though a useful approach for future research might be to choose or design bias tasks whose structural features are as similar as possible in order to permit more robust latent variables to be extracted, such design decisions might also allow method variance to contaminate the latent variables. It is also worth noting that the variability among our tasks is representative of the variability seen across studies purporting to measure implicit racial bias. If such variability affected the relations among our tasks, they also raise concerns about generalizing across studies using different tasks.

While the large number of tasks employed and the latent variable approach they permitted is a strength of the research, we must also consider the impact of fatigue associated with the completion of so many tasks. We tried to minimize this through frequent breaks, close monitoring for fatigue, and screening for below chance levels of responding, which was detected at a low rate. Nevertheless, we acknowledge that fatigue could have contributed noise to the data.

Despite the large number of tasks employed, inferences concerning the role of EFs in the expression of bias based on the current results are limited by the specific facets of EF we chose to measure. With perhaps one exception (Common EF and PDP automatic predicting WIT/FPST bias; see Figure 3) our models failed to account for the total variance in both bias expression and control during the tasks, suggesting that the bias tasks used here tap some other forms of control not represented in the unity and diversity model (see Sherman et al., 2008). Moreover, although latent variable models clearly permit better estimates of EF facets than do any single tasks (see Miyake et al., 2000) it is also clear that performance on some tasks relates more strongly to *individual differences* in underlying EF abilities than does performance on others. For example, despite its status as a prototypical inhibition measure (see Verbruggen & Logan, 2008), the stop-signal task was only modestly associated with the Common EF factor; likewise, the color-shape task was robustly associated with Shifting-specific ability but demonstrated a much smaller association with Common EF than did the other shifting tasks. Future research would benefit from both a broader conceptualization of EF abilities and better specificity in EF measurement.

Finally, we note that the goal of this research was to identify the qualitative nature of control implemented during implicit bias tasks rather than to document the conditions under which control operates (cf., Fujita, Trope, Cunningham, & Liberman, 2014; Gawronski & Bodenhausen, 2009; Sherman, 2006). Some definitions of control assume it requires intention and effort (Bargh, 1994), but researchers are increasingly recognizing that aligning behavior with goals can operate with little intention and/or effort (e.g., Fujita et al., 2014; Moskowitz et al., 1999). Our own data are broadly in line with this conclusion by showing that EF influences responses on implicit bias tasks when tight response deadlines limit the opportunity for reflective thought, but this issue requires additional research.

Conclusions

The current work represents a multidisciplinary union of theory and methods from contemporary cognitive science and social cognition brought to bear on a question that has intrigued scientists in both disciplines and that has considerable implications for our understanding of the meaning of implicit racial bias. Significant methodological advances include the relatively large number of participants recruited, the measurement of relevant constructs (EF and racial bias) with multiple representative tasks, the rigorous multisession procedure, the use of advanced quantitative modeling techniques, and the collection of data in three distinct geographic locations. In designing this project, our goal was to create a definitive dataset that could be used as a reference for the many scientists who are interested in core issues underlying racial bias and the impact of EF on implicit measures. Because these issues cross many disciplinary boundaries, our findings are of value not only to those directly interested in the role of EF on implicit bias, but also methodologists who seek to understand how implicit measures work, cognitive scientists who investigate the structure and functions of cognitive control, and social scientists of many orientations interested in race relations and the operation of implicit and explicit racial bias.

The current results make clear that various EFs do indeed relate to the racial bias expressed on implicit and explicit measures. Critically, these relationships depend greatly on the type of implicit measure used. In this way, our findings highlight not only the commonalities between various measures of racial bias, but also a number of intriguing differences among them. Although past research has documented some of the relationships shown here, past studies suffered from limited sample sizes, leaving open the possibility that their results were contaminated by both Type I and Type II errors. Our sample size offers some protection from these issues, and in this way serves as a roadmap for future research, highlighting areas where suggestive relationships need to be further examined, and where mysteries remain. Because of this, the current results stand as a statement on the field's current understanding of EF and implicit bias, highlighting both what we know about the implicit measurement of racial bias, and what requires further investigation.

References

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology, 91,* 652–661. http://dx.doi.org/10.1037/0022-3514.91.4.652

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neural signals for control. *Journal of Personality and Social Psychology, 94,* 60–74. http://dx.doi.org/10.1037/0022-3514.94.1.60

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality and Social Psychology, 84,* 738–753. http://dx.doi.org/10.1037/0022-3514.84.4.738

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science, 15,* 88–93. http://dx.doi.org/10.1111/j.0963-7214.2004.01502003.x

Amodio, D. M., & Mendoza, S. A. (2010). Implicit intergroup bias: Cognitive, affective, and motivational underpinnings. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition* (pp. 353–374). New York, NY: Guilford Press.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., pp. 1–40). Hillsdale, NJ: Erlbaum, Inc.

Bargh, J. A. (1999). The cognitive monster: The case against controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 361–382). New York, NY: Guilford Press.

Bartholow, B. D., Dickter, C. L., & Sestir, M. A. (2006). Stereotype activation and control of race bias: Cognitive control of inhibition and its impairment by alcohol. *Journal of Personality and Social Psychology, 90,* 272–287. http://dx.doi.org/10.1037/0022-3514.90.2.272

Bartholow, B. D., Henry, E. A., Lust, S. A., Saults, J. S., & Wood, P. K. (2012). Alcohol effects on performance monitoring and adjustment: Affect modulation and impairment of evaluative cognitive control. *Journal of Abnormal Psychology, 121,* 173–186. http://dx.doi.org/10.1037/a0023664

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science, 16,* 396–403. http://dx.doi.org/10.1111/j.1467-8721.2007.00534.x

Beer, J. S., Stallen, M., Lombardo, M. V., Gonsalkorale, K., Cunningham, W. A., & Sherman, J. W. (2008). The Quadruple Process model approach to examining the neural underpinnings of prejudice. *NeuroImage, 43,* 775–783. http://dx.doi.org/10.1016/j.neuroimage.2008.08.033

Blair, I. (2001). Implicit stereotypes and prejudice. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 359–374). Mahwah, NJ: Erlbaum.

Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the Implicit Association Test: Implications for criterion prediction. *Journal of Experimental Social Psychology, 42,* 192–212. http://dx.doi.org/10.1016/j.jesp.2005.07.003

Brauer, M., Wasel, W., & Niedenthal, P. (2000). Implicit and explicit components of prejudice. *Review of General Psychology, 4,* 79–101. http://dx.doi.org/10.1037/1089-2680.4.1.79

Brigham, J. C. (1993). College students' racial attitudes. *Journal of Applied Social Psychology, 23,* 1933–1967. http://dx.doi.org/10.1111/j.1559-1816.1993.tb01074.x

Burgess, P. W. (1997). Theory and methodology in executive function research. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 81–116). Hove, UK: Psychology Press.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology, 33,* 541–560. http://dx.doi.org/10.1006/jesp.1997.1329

Cohen, J., & Cohen, P. (1984). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). East Sussex, UK: Psychology Press.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82,* 407–428. http://dx.doi.org/10.1037/0033-295X.82.6.407

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The quad model of implicit task performance. *Journal of Personality and Social Psychology, 89,* 469–487. http://dx.doi.org/10.1037/0022-3514.89.4.469

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83,* 1314–1329. http://dx.doi.org/10.1037/0022-3514.83.6.1314

Correll, J., Park, B., & Smith, J. A. (2008). Colorblind and multicultural prejudice reduction strategies in high-conflict situations. *Group Processes & Intergroup Relations, 11,* 471–491. http://dx.doi.org/10.1177/1368430208095401

Correll, J., Urland, G. R., & Ito, T. A. (2006). Event-related potentials and the decision to shoot: The role of threat perception and cognitive control. *Journal of Experimental Social Psychology, 42,* 120–128. http://dx.doi.org/10.1016/j.jesp.2005.02.006

Crosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of Black and White discrimination and prejudice: A literature review. *Psychological Bulletin, 87,* 546–563. http://dx.doi.org/10.1037/0033-2909.87.3.546

Cunningham, W. A., Johnson, M. K., Raye, C. L., Chris Gatenby, J., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of Black and White faces. *Psychological Science, 15,* 806–813. http://dx.doi.org/10.1111/j.0956-7976.2004.00760.x

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12,* 163–170. http://dx.doi.org/10.1111/1467-9280.00328

De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology, 50,* 77–85.

De Houwer, J., & De Bruycker, E. (2007). The Implicit Association Test outperforms the Extrinsic Affective Simon Task as a measure of interindividual differences in attitudes. *British Journal of Social Psychology, 46,* 401–421.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin, 135,* 347–368. http://dx.doi.org/10.1037/a0014211

Denckla, M. B. (1996). A theory and model of executive function: A neuropsychological perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 263–278). Baltimore, MD: Brookes.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5–18. http://dx.doi.org/10.1037/0022-3514.56.1.5

Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton Trilogy revisited. *Personality and Social Psychology Bulletin, 21,* 1139–1150. http://dx.doi.org/10.1177/01461672952111002

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology, 82,* 835–848. http://dx.doi.org/10.1037/0022-3514.82.5.835

Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology, 22,* 22–37. http://dx.doi.org/10.1016/0022-1031(86)90039-9

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23,* 316–326. http://dx.doi.org/10.1177/0146167297233009

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology, 23,* 75–109. http://dx.doi.org/10.1016/S0065-2601(08)60318-4

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69,* 1013–1027. http://dx.doi.org/10.1037/0022-3514.69.6.1013

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition. research: Their meaning and use. *Annual Review of Psychology, 54,* 297–327. http://dx.doi.org/10.1146/annurev.psych.54.101601.145225

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology, 26,* 874–890. http://dx.doi.org/10.1080/13803390490510680

Friedman, N. P., Haberstick, B. C., Willcutt, E. G., Miyake, A., Young, S. E., . . . Hewitt, J. K. (2007). Greater attention problems during childhood predict poorer executive functioning in late adolescence. *Psychological Science, 18,* 893–900. http://dx.doi.org/10.1111/j.1467-9280.2007.01997.x

Friedman, N. P., Miyake, A., Altamirano, L. J., Corley, R. P., Young, S. E., . . . Hewitt, J. K. (2014). *Stability and change in executive function abilities from late adolescence to early adulthood: A longitudinal twin study*. Manuscript submitted for publication.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17,* 172–179. http://dx.doi.org/10.1111/j.1467-9280.2006.01681.x

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology, 47,* 1410–1430. http://dx.doi.org/10.1037/a0023750

Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General, 137,* 201–225. http://dx.doi.org/10.1037/0096-3445.137.2.201

Fujita, K. (2011). On conceptualizing self-control as more than the effortful inhibition of impulses. *Personality and Social Psychology Review, 15,* 352–366. http://dx.doi.org/10.1177/1088868311411165

Fujita, K., Trope, Y., Cunningham, W. A., & Liberman, N. (2014). What is control? A conceptual analysis. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 50–68). New York, NY: Guilford Press.

Gaertner, S. L., & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. *Social Psychology Quarterly, 46,* 23–30. http://dx.doi.org/10.2307/3033657

Gawronski, B., & Bodenhausen, G. V. (2009). Operating principles versus operating conditions in associative and propositional processes. *Behavioral and Brain Sciences, 32,* 207–208.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60,* 509–517. http://dx.doi.org/10.1037/0022-3514.60.4.509

Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice. *Journal of Experimental Social Psychology, 44,* 164–172. http://dx.doi.org/10.1016/j.jesp.2007.01.002

Gonsalkorale, K., Sherman, J. W., Allen, T. J., Klauer, K. C., & Amodio, D. M. (2011). Accounting for successful control of implicit racial bias: The roles of association activation, response monitoring, and overcoming bias. *Personality and Social Psychology Bulletin, 37,* 1534–1545. http://dx.doi.org/10.1177/0146167211414064

Govorun, O., & Payne, B. K. (2006). Ego depletion and prejudice: Separating automatic and controlled components. *Social Cognition, 24,* 111–136. http://dx.doi.org/10.1521/soco.2006.24.2.111

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27. http://dx.doi.org/10.1037/0033-295X.102.1.4

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464–1480. http://dx.doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197–216. http://dx.doi.org/10.1037/0022-3514.85.2.197

Hilgard, J., Bartholow, B. D., Dickter, C. L., & Blanton, H. (2014). Characterizing switching and congruency effects in the Implicit Association Test as reactive and proactive cognitive control. *Social Cognitive and Affective Neuroscience.* Advance online publication. http://dx.doi.org/10.1093/scan/nsu060

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31,* 1369–1385. http://dx.doi.org/10.1177/0146167205275613

Hofmann, W., Gschwendner, T., & Schmitt, M. (2005). On implicit-explicit consistency: The moderating role of individual differences in awareness and adjustment. *European Journal of Personality, 19,* 25–49. http://dx.doi.org/10.1002/per.537

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences, 16,* 174–180. http://dx.doi.org/10.1016/j.tics.2012.01.006

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424–453. http://dx.doi.org/10.1037/1082-989X.3.4.424

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in race categorization. *Psychological Science, 15,* 342–345. http://dx.doi.org/10.1111/j.0956-7976.2004.00680.x

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44,* 2017–2036. http://dx.doi.org/10.1016/j.neuropsychologia.2006.01.010

Huntsinger, J. R., Sinclair, S., & Clore, G. L. (2009). Affective regulation of implicitly measured stereotype and attitudes: Automatic and controlled components. *Journal of Experimental Social Psychology, 45,* 560–566. http://dx.doi.org/10.1016/j.jesp.2009.01.007

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513–541. http://dx.doi.org/10.1016/0749-596X(91)90025-F

Klauer, K. C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task-set switching account of the IAT effect. *Personality and Social Psychology Bulletin, 31,* 208–217. http://dx.doi.org/10.1177/0146167204271416

Klauer, K. C., Schmitz, F., Teige-Mocigemba, S., & Voss, A. (2010). Understanding the role of executive control in the implicit association test: Why flexible people have small IAT effects. *The Quarterly Journal of Experimental Psychology, 63,* 595–619. http://dx.doi.org/10.1080/17470210903076826

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the Implicit Association Test: A diffusion-model analysis. *Journal of Personality and Social Psychology, 93,* 353–368. http://dx.doi.org/10.1037/0022-3514.93.3.353

Klein, A., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika, 65,* 457–474. http://dx.doi.org/10.1007/BF02296338

Koole, S. L., Dijksterhuis, A., & van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology, 80,* 669–685. http://dx.doi.org/10.1037/0022-3514.80.4.669

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108,* 480–498. http://dx.doi.org/10.1037/0033-2909.108.3.480

Lambert, A. J., Payne, B. K., Shaffer, L. M., Jacoby, L. L., Chasteen, A., & Khan, S. (2003). Stereotypes as dominant responses: On the "social facilitation" of prejudice in anticipated public contexts. *Journal of Personality and Social Psychology, 84,* 277–295. http://dx.doi.org/10.1037/0022-3514.84.2.277

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21,* 59–80. http://dx.doi.org/10.1348/026151003321164627

Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67,* 808–817. http://dx.doi.org/10.1037/0022-3514.67.5.808

Maddux, W. W., Barden, J., Brewer, M. B., & Petty, R. E. (2005). Saying no to negativity: The effects of context and motivation to control prejudice on automatic evaluative responses. *Journal of Experimental Social Psychology, 41,* 19–35. http://dx.doi.org/10.1016/j.jesp.2004.05.002

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, *26*, 1124–1140. http://dx.doi.org/10.1037/0278-7393.26.5.1124

Mendoza, S. A., Gollwitzer, P. M., & Amodio, D. M. (2010). Reducing the expression of implicit stereotypes: Reflexive control through implementation intentions. *Personality and Social Psychology Bulletin, 36,* 512–523. http://dx.doi.org/10.1177/0146167210362789

Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the implicit association test. *Journal of Personality and Social Psychology, 85,* 1180–1192. http://dx.doi.org/10.1037/0022-3514.85.6.1180

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J.-C. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115,* 123–142. http://dx.doi.org/10.1016/j.actpsy.2003.12.004

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21,* 8–14. http://dx.doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100. http://dx.doi.org/10.1006/cogp.1999.0734

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice reduction efforts. *Journal of Personality and Social Psychology, 65,* 469–485. http://dx.doi.org/10.1037/0022-3514.65.3.469

Monteith, M. J., Ashburn-Nardo, L., Voils, C. I., & Czopp, A. M. (2002). Putting the brakes on prejudice: On the development and operation of cues for control. *Journal of Personality and Social Psychology, 83,* 1029–1050. http://dx.doi.org/10.1037/0022-3514.83.5.1029

Monteith, M. J., Lybarger, J. E., & Woodcock, A. (2009). Schooling the cognitive monster: The role of motivation in the regulation and control of prejudice. *Social and Personality Compass, 3,* 211–226. http://dx.doi.org/10.1111/j.1751-9004.2009.00177.x

Monteith, M. J., Sherman, J. W., & Devine, P. G. (1998). Suppression as a stereotype control strategy. *Personality and Social Psychology Review, 2,* 63–82. http://dx.doi.org/10.1207/s15327957pspr0201_4

Mooijaart, A., & Satorra, A. (2009). On insensitivity of the chi-square model test to nonlinear misspecification in structural equation models. *Psychometrika, 74,* 443–455. http://dx.doi.org/10.1007/s11336-009-9112-5

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81,* 111–121. http://dx.doi.org/10.1111/j.2044-8295.1990.tb02349.x

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology, 77,* 167–184. http://dx.doi.org/10.1037/0022-3514.77.1.167

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O’Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences, 15,* 453–459. http://dx.doi.org/10.1016/j.tics.2011.07.011

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user’s guide, 7th edition*. Los Angeles, CA: Muthén & Muthén.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General, 106,* 226–254. http://dx.doi.org/10.1037/0096-3445.106.3.226

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Oliver, M. B. (2003). African American men as “criminal and dangerous”: Implications of media portrayals of crime on the “criminalization” of African American men. *Journal of African American Studies, 7,* 3–18.

Park, S. H., Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice moderates the effect of cognitive depletion on unintended discrimination. *Social Cognition, 26,* 401–419. http://dx.doi.org/10.1521/soco.2008.26.4.401

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 81,* 181–192. http://dx.doi.org/10.1037/0022-3514.81.2.181

Payne, B. K. (2005). Conceptualizing control in social cognition: How executive control modulates the expression of automatic stereotyping. *Journal of Personality and Social Psychology, 89,* 488–503. http://dx.doi.org/10.1037/0022-3514.89.4.488

Payne, B. K., Burkley, M., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology, 94,* 16–31. http://dx.doi.org/10.1037/0022-3514.94.1.16

Payne, B. K., Lambert, A. J., & Jacoby, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology, 38,* 384–396. http://dx.doi.org/10.1016/S0022-1031(02)00006-9

Payne, B. K., Shimizu, Y., & Jacoby, L. L. (2005). Mental control and visual illusions: Toward explaining race-biased weapon misidentifications. *Journal of Experimental Social Psychology, 41,* 36–47. http://dx.doi.org/10.1016/j.jesp.2004.05.001

Phillips, L. H. (1997). Do “frontal tests” measure executive function? Issues of assessment and evidence from fluency tests. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 191–213). Hove, UK: Psychology Press.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75,* 811–832. http://dx.doi.org/10.1037/0022-3514.75.3.811

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Rabbitt, P. (1997). Introduction: Methodologies and models in the study of executive function. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 1–38). Hove, UK: Psychology Press.

Ranganath, K. A., Smith, C. T., & Nosek, B. A. (2008). Distinguishing automatic and controlled components of attitudes from direct and indirect measurement methods. *Journal of Experimental Social Psychology, 44,* 386–396. http://dx.doi.org/10.1016/j.jesp.2006.12.008

Richeson, J. A., Baird, A. A., Gordon, H. L., Heatherton, T. F., Wyland, C. L., . . . Shelton, J. N. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience, 6,* 1323–1328. http://dx.doi.org/10.1038/nn1156

Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science, 14,* 287–290. http://dx.doi.org/10.1111/1467-9280.03437

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General, 123,* 374–393. http://dx.doi.org/10.1037/0096-3445.123.4.374

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231. http://dx.doi.org/10.1037/0096-3445.124.2.207

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66,* 507–514. http://dx.doi.org/10.1007/BF02296192

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika, 75,* 243–248. http://dx.doi.org/10.1007/s11336-009-9135-y

Schlauch, R. C., Lang, A. R., Plant, E. A., Christensen, R., & Donohue, K. F. (2009). Effect of alcohol on race-biased responding: The moderating role of internal and external motivations to respond without prejudice. *Journal of Studies on Alcohol and Drugs, 70,* 328–336.

Schmitz, F., Teige-Mocigemba, S., Voss, A., & Klauer, K. C. (2013). When scoring algorithms matter: Effects of working memory load on different IAT scores. *British Journal of Social Psychology, 52,* 103–121. http://dx.doi.org/10.1111/j.2044-8309.2011.02057.x

Sherman, J. W. (2006). Clearing up some misconceptions about the Quad Model. *Psychological Inquiry, 17,* 269–276.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review, 115,* 314–335. http://dx.doi.org/10.1037/0033-295X.115.2.314

Sherman, J. W., Klauer, K. C., & Allen, T. J. (2010). Mathematical modeling of implicit social cognition: The machine in the ghost. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 156–175). New York, NY: Guilford Press.

Sherman, S. J., Rose, J. S., Koch, K., Presson, C. C., & Chassin, L. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. *Journal of Social and Clinical Psychology, 22,* 13–39. http://dx.doi.org/10.1521/jscp.22.1.13.22766

Siegel, E. F., Dougherty, M. R., & Huber, D. E. (2012). Manipulating the role of cognitive control while taking the implicit association test. *Journal of Experimental Social Psychology, 48,* 1057–1068. http://dx.doi.org/10.1016/j.jesp.2012.04.011

Stewart, B. D., & Payne, B. K. (2008). Bringing automatic stereotyping under control: Implementation intentions as efficient means of thought control. *Personality and Social Psychology Bulletin, 34,* 1332–1345. http://dx.doi.org/10.1177/0146167208321269

Stewart, B. D., von Hippel, W., & Radvansky, G. A. (2009). Age, race, and implicit prejudice: Using process dissociation to separate the underlying components. *Psychological Science, 20,* 164–168. http://dx.doi.org/10.1111/j.1467-9280.2009.02274.x

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662. http://dx.doi.org/10.1037/h0054651

Teuber, H.-L. (1972). Unity and diversity of frontal lobe functions. *Acta Neurobiologiae Experimentalis, 32,* 615–656.

van den Wildenberg, W. P. M., van Boxtel, G. J. M., van der Molen, M. W., Bosch, D. A., Speelman, J. D., & Brunia, C. H. M. (2006). Stimulation of the subthalamic region facilitates the selection and inhibition of motor responses in Parkinson's disease. *Journal of Cognitive Neuroscience, 18,* 626–636. http://dx.doi.org/10.1162/jocn.2006.18.4.626

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences, 12,* 418–424. http://dx.doi.org/10.1016/j.tics.2008.07.005

von Hippel, W. (2007). Aging, executive functioning, and social control. *Current Directions in Psychological Science, 16,* 240–244. http://dx.doi.org/10.1111/j.1467-8721.2007.00512.x

von Hippel, W., Silver, L. A., & Lynch, M. E. (2000). Stereotyping against your will: The role of inhibitory ability in stereotyping and prejudice among the elderly. *Personality and Social Psychology Bulletin, 26,* 523–532. http://dx.doi.org/10.1177/0146167200267001

Wen, Z., Marsh, H. W., & Hau, K.-T. (2010). Structural equation models of latent interactions: An appropriate standardized solution and its scale-free properties. *Structural Equation Modeling, 17,* 1–22. http://dx.doi.org/10.1080/10705510903438872

Wilcox, R. R., & Keselman, H. J. (2003). Modern robust data analysis methods: Measures of central tendency. *Psychological Methods, 8,* 254–274. http://dx.doi.org/10.1037/1082-989X.8.3.254

Willcutt, E. G., Pennington, B. F., Boada, R., Ogline, J. S., Tunick, R. A., . . . Olson, R. K. (2001). A comparison of the cognitive deficits in reading disability and attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology, 110,* 157–172. http://dx.doi.org/10.1037/0021-843X.110.1.157

Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors, 5,* 7–17.

Young, S. E., Friedman, N. P., Miyake, A., Willcutt, E. G., Corley, R. P., . . . Hewitt, J. K. (2009). Behavioral disinhibition: Liability for externalizing spectrum disorders and its genetic and environmental relation to response inhibition across adolescence. *Journal of Abnormal Psychology, 118,* 117–130. http://dx.doi.org/10.1037/a0014657

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151–175. http://dx.doi.org/10.1037/0003-066X.35.2.151

(*Appendix follows*)

Appendix

Additional Results

Table A1
Factor Scores and Loadings From Explicit Measures

	Factor 1: Personal attitude	Factor 2: Internal motivation	Factor 3: External motivation	Factor 4: Cultural stereotypes
Attitudes Toward Blacks (ATB)				
I would rather not have Black people live in the same apartment building I live in.	**.67**	.09	.13	−.01
I get very upset when I hear a White person make a prejudicial remark about Black people (R).	.34	**−.41**	.04	.02
Black and White people are inherently equal (R).	.32	−.10	−.04	.00
I would not mind at all if a Black family with about the same income and education as me moved in next door (R).	**.54**	.17	−.06	−.10
It would not bother me if my new roommate were Black (R).	**.63**	.17	.02	−.07
Whites should support Blacks in their struggle against discrimination and segregation (R).	**.57**	−.08	−.09	−.05
If a Black person were put in charge of me, I would not mind taking advice and direction from him or her (R).	**.62**	.08	−.03	−.04
I think that Black people look more similar to each other than White people do.	**.44**	−.02	.29	.06
The federal government should take decisive steps to override the injustices Black people suffer at the hands of local authorities (R).	.37	−.20	−.02	.01
I would probably feel somewhat self-conscious dancing with a Black person in a public place.	**.55**	.15	.12	.01
Interracial marriage should be discouraged to avoid the “who-am-I?” confusion which the children feel.	**.53**	.08	.00	−.06
Black people are demanding too much too fast in their push for equal rights.	**.48**	−.07	−.04	.00
I enjoy a funny racial joke, even if some people may find it offensive.	.01	**−.71**	.17	.10
If I had a chance to introduce Black visitors to my friends and neighbors, I would be pleased to do so (R).	**.67**	.16	−.04	−.08
I favor open housing laws that allow more racial integration of neighborhoods (R).	**.62**	.03	.01	−.08
Generally, Black people are not as smart as White people.	**.59**	.02	−.10	.02
Some Black people are so touchy about race that it is difficult to get along with them.	.34	−.28	.20	.14
It is likely that Black people will bring violence to neighborhoods when they move in.	**.66**	.09	−.02	.02
Racial integration (of schools, businesses, residences, etc.) has benefited both Whites and Black people (R).	.33	−.01	−.09	−.01
I worry that in the next few years I may be denied my application for a job or a promotion because of preferential treatment given to minority group members.	.32	−.19	.17	.01
Internal and External Motivation to Respond without Prejudice				
Because of today’s PC (politically correct) standards, I try to appear nonprejudiced toward Black people.	.04	−.09	**.74**	.01
I try to hide any negative thought about Black people in order to avoid negative reactions from others.	.13	.07	**.68**	−.01
I attempt to act in nonprejudiced ways toward Black people because it is personally important to me.	−.18	.32	.37	.03
I try to act nonprejudiced toward Black people because of pressure from others.	.18	−.23	**.74**	−.06
According to my personal values, using stereotypes about Black people is ok (R).	−.30	**.40**	−.09	.01
If I acted prejudiced toward Black people, I would be concerned that others would be angry with me.	−.05	.02	**.67**	−.01
Being nonprejudiced toward Black people is important to my self-concept.	−.32	.20	.34	.00
I attempt to appear nonprejudiced toward Black people in order to avoid disapproval from others.	.16	−.22	**.82**	−.03
Because of my personal values, I believe that using stereotypes about Black people is wrong.	−.32	**.40**	.13	.03
I am personally motivated by my beliefs to be nonprejudiced toward Black people.	−.37	.27	.28	.02

(*Appendix continues*)

Table A1 (*continued*)

	Factor 1: Personal attitude	Factor 2: Internal motivation	Factor 3: External motivation	Factor 4: Cultural stereotypes
Nonscale measures				
Feeling thermometer bias	**.57**	.02	.21	.00
Personal stereotype bias —Aggressive	**.51**	.10	−.08	.22
Personal stereotype bias —Violent	**.59**	.12	.03	.15
Personal stereotype bias—Dangerous	**.52**	.19	.09	.09
Perceived cultural stereotype bias—Aggressive	−.06	−.03	−.04	**.83**
Perceived cultural stereotype bias—Violent	.00	.00	−.01	**.93**
Perceived cultural stereotype bias—Dangerous	.02	.07	−.02	**.85**
Funding allocation bias	.39	−.05	.07	−.04
Motivation to Control Prejudiced Reactions (MCPR)				
In today's society it is important that one not be perceived as prejudiced in any manner.	−.17	.07	.28	.02
I always express my thoughts and feelings, regardless of how controversial they might be (R).	.20	**.41**	.08	.04
I get angry with myself when I have a thought or feeling that might be considered prejudiced.	−.10	**.45**	.20	−.04
If I were participating in a class discussion and a Black student expressed an opinion with which I disagreed, I would be hesitant to express my own viewpoint.	.27	.27	.07	−.10
Going through life worrying about whether you might offend someone is just more trouble than it's worth (R).	.05	**.48**	−.01	.00
It's important to me that other people not think I'm prejudiced.	−.11	.26	**.40**	−.04
I feel it's important to behave according to society's standards.	.18	.21	.29	.07
I'm careful not to offend my friends, but I don't worry about offending people I don't know or don't like (R).	−.16	.31	−.16	.06
I think that it is important to speak one's mind rather than to worry about offending someone (R).	.26	**.66**	−.11	.05
It's never acceptable to express one's prejudices.	−.08	.36	.00	.00
I feel guilty when I have a negative thought or feeling about a Black person.	−.06	.36	.34	−.06
When speaking to a Black person, it's important to me that he/she not think I'm prejudiced.	.00	.18	**.51**	−.01
It bothers me a great deal when I think I've offended someone, so I'm always careful to consider other people's feelings.	.05	**.61**	.09	.08
If I have a prejudiced thought or feeling, I keep it to myself.	−.08	**.53**	.12	.07
I would never tell jokes that might offend others.	.16	**.81**	−.12	−.03
I'm not afraid to tell others what I think, even when I know they disagree with me (R).	.32	**.46**	−.07	.00
If someone who made me uncomfortable sat next to me on a bus, I would not hesitate to move to another seat (R).	.04	.12	.04	.06

Note. Item loadings .40 and greater are bolded. R indicates a reserve scored item.

(*Appendix continues*)

Table A2

Intercorrelations Among Factor Scores and Individual Explicit Measures

	Extracted factor			
Measure	Personal attitude	Internal motivation	External motivation	Cultural stereotype
Personal attitude (Factor 1)	—			
Internal motivation (Factor 2)	−.58*	—		
External motivation (Factor 3)	−.22*	.61*	—	
Cultural stereotypes (Factor 4)	.03	−.05	.01	—
Attitudes toward Blacks	.94*	−.58*	−.20*	.03
Internal motivation to control prejudice	−.66*	.79*	.61*	.00
External motivation to control prejudice	−.01	.37*	.91*	.00
Motivation to control prejudiced reactions	−.38*	.91*	.68*	−.02
Feeling thermometer bias	.51*	−.17*	.10*	.01
Personal stereotype bias	.59*	−.18*	−.03	.20*
Perceived cultural stereotype bias	.00	−.04	.00	.98*
Funding allocation bias	.41*	−.22*	−.04	−.03

Note. Higher values indicate more negative Personal Attitudes, higher internal and external motivation to control prejudice, and greater stereotyping of Blacks. Correlations are maximum likelihood estimates (from Mplus) based on all data, adjusted for missing observations. Total $N = 481$.

* $p \leq .05$.

Table A3

Mean Accuracy Rates and Reaction Times for Implicit Tasks as a Function of Trial/Block Type

Trial/Block type	Accuracy	RT
Weapons Identification Task		
Black—Gun	.80 (.10)	389 (29)
Black—Tool	.68 (.14)	405 (32)
White—Gun	.72 (.12)	394 (30)
White—Tool	.77 (.12)	397 (30)
First Person Shooter Task		
Black—Armed	.78 (.11)	463 (24)
Black—Unarmed	.72 (.13)	494 (23)
White—Armed	.74 (.09)	460 (25)
White—Unarmed	.70 (.12)	492 (23)
Implicit Association Test		
Stereotype congruent	.94 (.04)	645 (102)
Stereotype incongruent	.91 (.07)	805 (174)

Note. RT = reaction time. Standard deviations in parentheses.

Received October 8, 2013
Revision received November 3, 2014
Accepted November 7, 2014 ■