



Dissociating automatic associations: Comparing two implicit measurements of race bias

Hannah I. Volpert-Esmond¹ | Laura D. Scherer² | Bruce D. Bartholow¹

¹Department of Psychological Sciences, University of Missouri, Columbia, MO, USA

²Division of Cardiology, University of Colorado School of Medicine, Denver, CO, USA

Correspondence

Hannah I. Volpert-Esmond, Department of Psychological Sciences, 210 McAlester Hall, University of Missouri, Columbia, MO 65211, USA.
Email: hannah.volpert@mail.missouri.edu

Funding information

National Science Foundation, Grant/Award Number: 1529788; National Institute on Alcohol Abuse and Alcoholism, Grant/Award Number: AA020970

Abstract

Weak correspondence across different implicit bias tasks may arise from the contribution of unique forms of automatic and controlled processes to response behavior. Here, we examined the correspondence between estimates of automatic and controlled processing derived from two sequential priming tasks with identical structure and timing designed to separately measure stereotypic (Weapons Identification Task; WIT) and evaluative (Affective Priming Task; APT) associations. Across two studies using predominantly White samples, three consistent patterns emerged in the data: (a) stereotypic bias was stronger for Black targets, whereas evaluative bias was stronger for White targets; (b) overall response accuracy bias correlated modestly across the two tasks; and (c) multinomial processing tree estimates of controlled processing corresponded much more strongly than estimates of automatic processing. These findings support models positing distinct learning and memory systems for different forms of race bias, and suggest that these differing forms contribute to estimates of automatic associations.

KEYWORDS

implicit race bias, process dissociation procedure, multiple memory systems, Weapons Identification task, affective priming

1 | INTRODUCTION

The use of indirect or “implicit” measures to investigate intergroup attitudes has become commonplace. Measures such as the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), the Weapons Identification Task (WIT; Payne, 2001), and evaluative priming procedures (Fazio, Jackson, Dunton, & Williams, 1995) are valued because of their promise to measure attitudes that people are unwilling or unable to report (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). This is particularly valuable for the study of racial bias, where explicitly reported attitudes have become more egalitarian but biased behavior persists (Dovidio, Gaertner, & Pearson, 2017). In some studies, implicit racial bias has been connected to intergroup behaviors as subtle as smiling and making eye contact (Dovidio, Kawakami, & Gaertner, 2002) and as severe as the

disproportionate use of lethal force by police on African-American men (Hehman, Flake, & Calanchini, 2017). Importantly, however, debate continues on how and when performance in implicit bias tasks is related to real-world behavior (e.g., Greenwald, Banaji, & Nosek, 2015; Hall et al., 2015; Kurdi et al., 2019; Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013) and precisely what implicit bias represents (Brownstein, Madva, & Gawronski, 2019; De Houwer, 2019; Greenwald & Banaji, 2017; Machery, 2017; Payne, Vuletich, & Lundberg, 2017).

A persistent problem with the measurement of implicit bias is that correlations among different implicit bias tasks are often low (Amodio & Devine, 2006; Ito et al., 2015). In part, low correspondence between tasks has been attributed to their poor psychometric properties, such as low reliability (Banse, Seise, & Zerbes, 2001; De Houwer et al., 2009; Kawakami & Dovidio, 2001). However,

conceptual issues pertaining to the nature of implicit bias as a construct also contribute to low correspondence across different tasks. For example, tasks differ with regard to the specific forms of bias they measure (e.g., stereotypic vs. evaluative; Amodio & Devine, 2006; Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Jacoby, 1991; Payne, 2001). In addition, responses on measures of implicit bias are known to be influenced by a number of control-related processes (Conrey et al., 2005; Jacoby, 1991; Payne, 2001), and to engage those processes to differing degrees (e.g., Ito et al., 2015; Klauer, Schmitz, Teige-Mocigemba, & Voss, 2010). To date, researchers have not given full consideration to how the type of automatic or controlled processes that are engaged in a particular task contributes to low intercorrelations across tasks. The purpose of the current studies was to investigate how the type of automatic associations measured in a given task (evaluative vs. stereotypic associations with race) affects performance and correspondence between estimates of automatic and controlled processing across two sequential priming tasks used to measure implicit racial bias.

1.1 | Contributions of automatic and controlled processes

Attitude researchers have long understood the utility in separating the influence of automatic or reflexive and controlled or reflective processes on behavior (see Chaiken & Trope, 1999; Smith & DeCoster, 2000; Strack & Deutsch, 2004). In recent years, this interest has led to the development of multinomial “processing tree” models meant to separately estimate the influence of different kinds of underlying processes on response behavior in implicit bias tasks. Processing tree models use patterns of response accuracy on different types of trials to mathematically separate theoretically distinct processes that jointly contribute to responses in a particular task (Hütter & Klauer, 2016). Several such models have been proposed (e.g., Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Payne & Bishara, 2009; Stahl & Degner, 2007), all of which aim to separately estimate the contribution of automatic and non-automatic processes to behavior in implicit bias tasks. Importantly, such estimates are atheoretical, and model parameters can only be related to psychological processes through empirical work that establishes the validity of the theoretical interpretations of the parameters (Hütter & Klauer, 2016).

One well-established model, the Process Dissociation Procedure (Payne, 2001, 2005), estimates two parameters thought to represent the contribution of controlled (PDP-Control) and automatic (PDP-Auto) processes on binary classification of targets (see Jacoby, 1991). In the traditional conception of the PDP, each parameter estimate represents the *magnitude* of the influence of controlled or automatic processing on response behavior. Validating studies using the PDP with implicit bias tasks have shown that PDP-Control estimates decrease with shorter response deadlines, suggesting decreased ability to control responses when forced to respond quickly (Payne, 2001), and are related to the strength of executive functioning abilities (Ito

et al., 2015; Payne, 2005). PDP-Control estimates also decrease as a function of age due to decreases in inhibitory ability, resulting in the expression of greater bias in older adults (Stewart, von Hippel, & Radvansky, 2009). PDP-Auto estimates are larger when participants are encouraged to use the race of the prime to identify weapons, similarly to racial profiling (Payne, Lambert, & Jacoby, 2002), and increase when White participants anticipated interacting with a Black partner compared with a White partner, interpreted as evidence that anxiety when anticipating an interracial interaction increases the strength of automatic processing (Amodio & Hamilton, 2012). These and other studies that have investigated the influence of experimental manipulations or correspondence with other theoretical relevant constructs contribute to the validity of PDP estimates as measurements of controlled and automatic processes related to implicit bias.

However, because different psychological processes may be engaged in different tasks or contexts, multinomial modeling tree parameters may index different controlled or automatic processes depending on the task or context. This is particularly true when considering tasks intended to measure different forms of bias, which stem from different types of learned associations. The Multiple Memory Systems Model of social cognition (Amodio, 2019; Amodio & Ratner, 2011) underscores that distinct types of implicit associations (e.g., semantic vs. evaluative) arise from different kinds of learning experiences, involve distinct neurophysiological circuits, and are expressed in differing behavioral responses. For instance, whereas semantic associations reflect verbal learning experiences (McClelland & Rumelhart, 1985; Sloman, 1996) rooted mainly in anterior temporal cortex (see Martin, 2007; Ralph, Jefferies, Patterson, & Rogers, 2017), evaluative associations can reflect both verbal learning and aversive conditioning mediated mainly by the central nucleus of the amygdala (LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998; LeDoux, 2000).

From a methodological standpoint, the type of automatic association relevant to performance in a given task depends on the type of stimuli respondents are asked to categorize. In sequential priming tasks, for example, race-related primes (e.g., faces varying by race) arguably activate both semantic (i.e., stereotypes) and evaluative (i.e., prejudice-related) associations, among others. It stands to reason that differences in the automatic processes engaged in a semantic/stereotypic versus evaluative bias task will be reflected in the correspondence between model parameters estimating the contribution of automatic processing in each task. If subsequent targets represent content semantically related to race categories (e.g., handguns represent the American stereotype that young Black men are violent; see Correll, Park, Judd, & Wittenbrink, 2002; Payne, 2001), then task performance primarily will be driven by stereotypic associations. In contrast, if target stimuli represent evaluative content (e.g., words describing positive or negative constructs; see Fazio et al., 1995), then task performance primarily will be driven by evaluative associations. Parameter estimates of automatic processes modeled from behavior in these different tasks would then theoretically reflect different forms of automatic associations, and the degree of correspondence in parameter estimates derived from

the two tasks should reflect (dis)similarity between the underlying automatic associations.

Recent research by Calanchini, Sherman, Klauer, and Lai (2014) supports this idea. These authors had participants complete pairs of IATs that varied in their conceptual overlap, operationalized as the extent to which the tests used similar attribute categories (e.g., pleasant and unpleasant words) and target categories (e.g., race). As predicted, the cross-task correlation of automatic associations derived from the quadruple process model (Conrey et al., 2005), which are virtually equivalent to PDP-Auto estimates (Payne & Bishara, 2009), varied systematically according to the degree of conceptual overlap in the various IATs ($r_s = .50-.58$ across tasks with high overlap; $r_s = .20-.36$ across tasks with low overlap).

Model parameters reflecting control-related processes are similarly expected to correspond as a function of similarity of the control-related processes engaged in a particular task. In a recent example, Ito et al. (2015) had participants complete three implicit bias tasks designed to measure stereotypic associations between Black men and violence/danger—the First Person Shooter Task (FPST; Correll, Park, Judd, & Wittenbrink, 2002), the WIT, and the IAT. Individual differences in control-related abilities also were assessed using nine tasks that measure different facets of executive functioning (EF; Friedman & Miyake, 2017; Miyake & Friedman, 2012). Ito et al. found that performance on both the WIT and the FPST was negatively associated with a “Common EF” factor that reflects variability common to all nine EF tasks (see Miyake & Friedman, 2012), suggesting the WIT and FPST engage similar forms of control. Consistent with the idea that correspondence in model estimates reflects correspondence in the type of controlled processing engaged, control estimates from the PDP were highly correlated between the WIT and FPST ($r = .61$). In contrast, performance on the IAT was associated with a facet of executive functioning related to shifting or switching attention (also see Klauer et al., 2010) but not with Common EF. As a result, control estimates from the IAT showed more modest correspondence with estimates derived from the Shooter Task ($r = .29$) and WIT ($r = .37$), although in a separate study, control estimates from various IATs with identical structure but different stimuli correlated very highly with each other ($r_s = .67-.82$; Calanchini, Sherman, Klauer, & Lai, 2014). Thus, control estimates from tasks that engage different forms of executive functioning or cognitive control correlate to a smaller degree than control estimates from tasks that engage the same form of executive functioning.

Together, these previous results suggest that correspondence in model estimates of controlled and automatic processing across tasks reflects similarity in the *type* of automatic or controlled processing being engaged in each task. To date, however, no previous research has used a multinomial processing tree approach to directly compare the magnitude of association in automatic and control estimates across sequential priming tasks designed to measure race-related stereotypic versus evaluative associations. Although this issue has been examined by comparing automatic and controlled estimates across different versions of the IAT

(Calanchini et al., 2014), examining how automatic associations influence response behavior in sequential priming tasks is important for several reasons. First, as just reviewed, the IAT and sequential priming tasks engage different control-related processes (Ito et al., 2015; Klauer et al., 2010). Given that control-related processes constrain the expression of automatic associations (Payne & Bishara, 2009), differences in the type of controlled process engaged by different tasks may alter the degree to which automatic processing influences response behavior. Second, the IAT and sequential priming tasks rely on different mechanisms through which automatic associations with race affect responses. Sequential priming effects are theorized to occur either through semantic priming (see Cameron et al., 2012; Gawronski & Hahn, 2019; Wentura & Degner, 2010) or through response priming (see Bartholow, Riordan, Saults, & Lust, 2009; Klinger, Burton, & Pitts, 2000). In contrast, the mechanism by which automatic associations affect responses on the IAT is less well understood (Teige-Mocigemba, Klauer, & Sherman, 2010); stimulus–response compatibilities, a random walk model, and task switching, among others, have been proposed as potential mechanisms (Brendl, Markman, & Messner, 2001; De Houwer, 2001; Klauer & Mierke, 2005; Teige-Mocigemba et al., 2010).

Lastly, compared to the IAT, sequential priming tasks are less susceptible to extrapersonal information unrelated to individual attitudes (Degner & Wentura, 2010; Han, Olson, & Fazio, 2006); the IAT has been proposed to measure societal views that do not necessarily correspond to personal attitudes (Olson & Fazio, 2004; Payne et al., 2017). Given these unique properties, comparison of automatic and controlled processing across sequential priming tasks that assess evaluative versus stereotypic associations is important for understanding whether previous findings based on the IAT (e.g., Amodio & Devine, 2006; Calanchini et al., 2014) generalize beyond the constraints imposed by that measure.

In the current studies, participants completed a WIT (Payne, 2001) to assess stereotypic associations between White vs. Black targets and armed violence, and an affective priming task (APT; Fazio et al., 1995) to assess associations between White vs. Black targets and evaluative constructs. To control for differences in the tasks' structural features and the types of control-related processes they engaged (Ito et al., 2015; Klauer et al., 2010; Payne, 2005), both tasks used a sequential priming trial structure in which a race prime (face) preceded a target (object or word) that participants had to classify prior to a 500-ms response deadline. Given that (a) tasks that engage the same type of executive functioning ability correspond highly in terms of controlled processing estimates from the PDP or other multinomial processing trees and (b) that both the WIT and the APT are sequential priming tasks associated with the “Common EF” facet of executive function (Friedman & Miyake, 2017; Ito et al., 2015; Miyake & Friedman, 2012), we predicted strong correspondence in PDP-Control estimates across the two sequential priming tasks. In contrast, because of the differences in the types of automatic processing engaged in each task, PDP-Auto estimates were

expected to correspond weakly (Calanchini et al., 2014). To ensure that any observed associations are reliable (see Stanley & Spence, 2014) and to better characterize the magnitude of observed effects (see Spellman, 2013), we conducted two experiments using the same measures and procedures¹ and separate samples of participants drawn from the same population.

2 | METHOD

2.1 | Participants

2.1.1 | Study 1

One hundred-one undergraduates (30 men, 70 women, 1 unidentified) from Introductory Psychology courses at a large, Midwestern university participated for credit toward a course requirement. Eighty-nine self-identified as White, one as American Indian/Alaska Native, four as Asian, four as Black, and one as of more than one race. Two did not indicate their race. None identified as Latino/a. With this sample size, we had 80% power to detect correlations $> .24$.

2.1.2 | Study 2

An important aim of Study 2 was to characterize the magnitude of effects using a larger sample, thereby supporting increased reliability of the estimates by reducing error variance (see Cocchetti, 1999; Stanley & Spence, 2014). Thus, the sample size from Study 1 was roughly doubled for Study 2. Two hundred and six undergraduates (109 women, 94 men, 3 unidentified) from Introductory Psychology courses participated for credit toward a course requirement. One hundred and sixty-six self-identified as White, eleven as Asian, fifteen as Black, ten as of more than one race, and four did not respond. Nine identified as Hispanic. With this sample size, we had 80% power to detect correlations $> .17$.

2.2 | Measures and materials

Internal and external motivation to be unbiased (Plant & Devine, 1998) was self-reported but results are not discussed. See Appendix S1 for more information.

2.2.1 | Weapons Identification Task

Stimuli and instructions were adapted from Payne (2001). During each trial, a 1000-ms fixation cross preceded a prime (Black or

White face; 200 ms), followed immediately by a target (gun or tool; 200 ms), which was then covered by a visual mask (300 ms). Prime stimuli consisted of 24 black-and-white photographs of White and Black men's faces cropped to exclude peripheral features (e.g., hair, clothes). Target stimuli consisted of four black-and-white photographs of handguns and four black-and-white photographs of hand tools (e.g., pliers, crescent wrench). Participants were asked to categorize the target on each trial as a gun or a tool using one of two response keys (counterbalanced across participants) within 500 ms following target onset. Responses made after the deadline elicited a "TOO SLOW!" message, displayed in red text for 500 ms. The ITI was randomly jittered across trials (800, 1000, or 1200 ms). Participants completed 16 practice trials followed by 192 experimental trials, comprising 48 trials of each prime-target pairing (i.e., Black-tool, Black-gun, White-tool, and White-gun) presented in a randomized order.

2.2.2 | Affective Priming Task

Like the WIT, trials in the APT consist of a prime stimulus followed by a target that participants must classify as quickly as possible. In the version used here, the prime stimuli, trial structure, and timing parameters, including the use of a response deadline, were identical to those used in the WIT. Although some APTs used to investigate implicit racial bias have not used a response deadline (e.g., Fazio et al., 1995; Livingston & Brewer, 2002), use of response windows or deadlines and measurement of accuracy rather than reaction time as the dependent variable in the APT have been advocated (a) to prevent possible dilution of priming effects, and (b) to reduce the use of faking strategies (e.g., Degner, 2009). The target stimuli were eight positive nouns (*love, vacation, joy, romance, paradise, success, beauty, and smile*) and eight negative nouns (*garbage, vomit, poison, sewage, pest, despair, cockroach, and disgust*) with no race-stereotypic content (Livingston & Brewer, 2002). Participants completed 16 practice trials, followed by 192 experimental trials, comprising 48 trials of each prime-target pairing (i.e., Black-positive, Black-negative, White-positive, and White-negative) presented in a randomized order.

2.3 | Procedure

The order in which the tasks were completed was randomized across participants. In both experiments, participants were randomly assigned using a counterbalancing procedure to one of two between-subjects conditions in which they completed the WIT and APT either in the presence of an experimenter (i.e., *Observer Present*) or alone (i.e., *Observer Absent*). In the *Observer Absent* condition, after explaining the tasks, the experimenter said, "I'm going to step out of the room during the tasks so you don't get distracted. I'll be right outside the door, so if you have any questions during the instructions or after the practice trials, feel free to come ask

¹A minor procedural difference between the studies was that, in Study 2, participants responded to several questions related to task-related state anxiety, frustration, confidence, attention, and effort (see Appendix S1). Participants responded to the items four times over the course of the experiment—at the midpoint and end of each task.

me." In the *Observer Present* condition, the experimenter remained in the room, saying, "Feel free to ask me any questions after the practice trials if you're confused by the instructions. Otherwise, just ignore me and concentrate on the tasks." The experimenter held a clipboard and sat in a chair to the left and slightly behind the participant. From this vantage point, the experimenter was able to see the participant (and their hands providing behavioral responses) as well as the computer screen. The participant was also able to see the experimenter in their peripheral vision while completing the tasks.

The presence versus absence of an observer did not significantly affect task performance or estimates of automatic or controlled processes in either experiment, possibly because the structure of the tasks imposed constraints on controlled processing that were not further affected by the presence of the observer. All analyses including effects of this manipulation are reported in Appendix S1.

2.4 | Data analysis

2.4.1 | Exclusions: Study 1

Data from one participant were lost for both tasks due to a computer malfunction. Data from three other participants (one in the APT and two in the WIT) were excluded for not following instructions (e.g., falling asleep, mixing up the buttons, using one response button >85% of the time). This left 98 participants with WIT data, 99 with APT data, and 97 with usable data for both tasks.²

2.4.2 | Exclusions: Study 2

APT data from two participants were lost due to computer malfunction. APT data from three additional participants were rejected for not following instructions (using one response button >85% of the time). WIT data from one participant were lost due to computer malfunction, and data from another participant was rejected for not following instructions (using one response button >85% of the time). This left 204 participants with WIT data, 201 with APT data, and 201 with usable data for both tasks.

²Additional analyses were conducted that excluded non-native English speakers and participants whose accuracy was >2 SDs below the mean (thresholds of 36% and 33% accuracy in the WIT and APT in Study 1, respectively, and 34% and 33% in the WIT and APT in Study 2, respectively). According to these criteria, Study 1 analyses included 92 participants' and 93 participants' WIT and APT data, respectively. Study 2 analyses included 199 participants' and 198 participants' WIT and APT data, respectively. However, even with these exclusions, the pattern of results remained the same. The only minor change was a decrease in the correlation between response accuracy bias scores across tasks, decreasing from $\beta = 0.23, p = .022$ to $\beta = .19, p = .081$. All other slope or interaction estimates remained very similar, including the difference in the degree to which PDP-Auto and PDP-Control estimates corresponded between tasks. Thus, we present analyses including all participants except those not following instructions to keep as large of a sample size as possible.

2.4.3 | Analytic approach

Analyses within each task included all usable data for that task. Regressions comparing the two tasks included only individuals with data for both tasks; however, mixed models do not require listwise deletion and thus included all participants with data for at least one task ($n = 101$ in Study 1; $n = 204$ in Study 2). To facilitate comparison of results across tasks using similar language, the 2 (Race) \times 2 (Target) structure of both tasks is described in the analyses in terms of bias-congruency. Specifically, the predicted Prime \times Target interactions for both tasks are presented as an effect of *congruency*, representing the extent to which the primes facilitated responses to stereotypically (WIT) or evaluatively (APT) congruent targets (i.e., Black primes facilitating categorization of guns and negative words; White primes facilitating categorization of tools and positive words). Analyses conducted with multilevel models rather than repeated measures ANOVA are reported in Appendix S1; these analyses produced patterns of effects identical to those reported here. As stimulus could not be included as a random factor in either set of models, the presented results are limited to the specific stimuli used here.

2.4.4 | Calculation of PDP estimates

PDP estimates were first created for Black-prime and White-prime trials separately (see Payne, 2001) and are given in Table 1.³ To create task-wide automatic estimates (PDP-Auto) for comparison across tasks, White-prime *Auto* estimates were partialled out from Black-prime *Auto* estimates (see Payne, 2005). This residual represents the influence of Black-related automatic processing while accounting for the influence of White-related automatic processing.⁴ Task-wide control estimates (PDP-Control) were calculated as the mean of *Control* estimates from Black and White trials, reflecting the argument that controlled processes are similar for all trials, regardless of prime race (see Ito et al., 2015; Payne, 2005).

Data and analyses can be found at <https://github.com/hiv8r3/ComparingBias-project>. Within this manuscript and Appendix S1, all measures, conditions, and data exclusions have been reported, including analyses regarding self-reported motivation to be unbiased that are reported in Appendix S1. All methods were approved by the University of Missouri IRB, and all participants were treated in accordance with ethical principles as laid out by the American Psychological Association.

³Control (C) estimates were calculated by subtracting the probability of making an error on a bias-incongruent trial (e.g., a Black face followed by a tool in the WIT or a positive word in the APT) from the probability of being correct on a bias-congruent trial (e.g., a Black face followed by a gun in the WIT or a negative word in the APT). Automatic (A) estimates were calculated by dividing the probability of making an error on an incongruent trial by $(1-C)$. Trials in which the participant did not respond before the deadline were excluded from the calculation, as responses given after the deadline were not recorded. Additionally, any negative estimates were replaced with 0.

⁴Parallel analyses using difference scores (Amodio & Hamilton, 2012; Ito et al., 2015) revealed a similar pattern of results (see Appendix S1).

| | WIT | | APT | |
|-------------|--------------|--------------|--------------|--------------|
| | Black primes | White primes | Black primes | White primes |
| Study 1 | | | | |
| PDP-Control | 0.41 (0.27) | 0.42 (0.23) | 0.40 (0.25) | 0.42 (0.23) |
| PDP-Auto | 0.62 (0.17) | 0.59 (0.18) | 0.57 (0.16) | 0.58 (0.19) |
| Study 2 | | | | |
| PDP-Control | 0.45 (0.27) | 0.45 (0.26) | 0.40 (0.26) | 0.39 (0.27) |
| PDP-Auto | 0.60 (0.18) | 0.57 (0.18) | 0.53 (0.18) | 0.62 (0.17) |

Note: Numbers in parentheses are standard deviations. PDP-Control and PDP-Auto = controlled and automatic processing estimates, respectively, derived from the process dissociation procedure. Abbreviations: APT, Affective Priming Task; WIT, Weapons Identification Task.

3 | RESULTS

For each study, we first present preliminary analyses that test for the presence of a Prime \times Target interaction representing biased responding, separately in each task. Once the presence of bias is confirmed in each task, we then compare behavioral responses across the tasks in three ways: (a) We test the 3-way Prime \times Target \times Task interaction predicting error rates in each trial type; (b) We examine the correlation between response accuracy bias scores between tasks, and (c) We compare multinomial processing tree estimates from the PDP across tasks. Last, we corroborate results obtained from the PDP with the ABC model (Stahl & Degner, 2007), an alternative multinomial processing tree model.

3.1 | Study 1

3.1.1 | Preliminary analyses

Mean error rates (and *SDs*) for each condition of the WIT and APT are given in Table 2 and displayed in Figure 1. Reaction times are also reported in Table 2 but were not analyzed. A repeated measures analysis of variance (rANOVA) on the error rates from the WIT showed the predicted effect of congruency (i.e., Prime \times Target interaction), $F(1, 97) = 91.4, p < .001, \eta^2_p = 0.48$ (90% CI: [0.37, 0.57]). Guns were categorized more accurately than tools following Black faces, $F(1, 97) = 102.4, p < .001, \eta^2_p = 0.51$, (90% CI: [0.40, 0.60]), whereas tools were categorized more accurately than guns following White faces, $F(1, 97) = 8.54, p = .004, \eta^2_p = 0.08$ (90% CI: [0.02, 0.18]). Similarly, rANOVA on the error rates from the APT also showed a predicted effect of congruency (i.e., Prime \times Target interaction), $F(1, 98) = 40.64, p < .001, \eta^2_p = 0.29$, (90% CI: [0.17, 0.40]). Negative words were categorized more accurately than positive words following Black faces, $F(1, 98) = 5.72, p = .019, \eta^2_p = 0.05$, (90% CI: [0.00, 0.14]), and positive words were categorized more accurately than negative words following White faces, $F(1, 98) = 41.15, p < .001, \eta^2_p = 0.30$, (90% CI: [0.17, 0.40]). These interactions confirmed

TABLE 1 Mean control and automatic estimates (and *SDs*) for Black-prime and White-prime Trials in the WIT and APT

TABLE 2 Mean error rates (and *SDs*) and reaction times (and *SDs*) as a function of prime and target in the WIT and APT

| | Black primes | White primes |
|----------|-----------------------|-----------------------|
| Study 1 | | |
| WIT | | |
| Gun | 0.31 (0.15)//373 (74) | 0.41 (0.13)//381 (79) |
| Tool | 0.47 (0.15)//398 (79) | 0.36 (0.15)//387 (75) |
| APT | | |
| Negative | 0.39 (0.16)//383 (80) | 0.47 (0.16)//392 (84) |
| Positive | 0.43 (0.15)//384 (84) | 0.34 (0.15)//379 (78) |
| Study 2 | | |
| WIT | | |
| Gun | 0.31 (0.18)//368 (81) | 0.39 (0.18)//374 (85) |
| Tool | 0.44 (0.17)//386 (91) | 0.35 (0.17)//376 (86) |
| APT | | |
| Negative | 0.42 (0.17)//374 (93) | 0.49 (0.17)//383 (95) |
| Positive | 0.42 (0.15)//370 (98) | 0.34 (0.16)//363 (92) |

Note: Numbers in parentheses are standard deviations. Numbers to the right of forward slashes are mean (and *SD*) reaction times in milliseconds (correct response trials only).

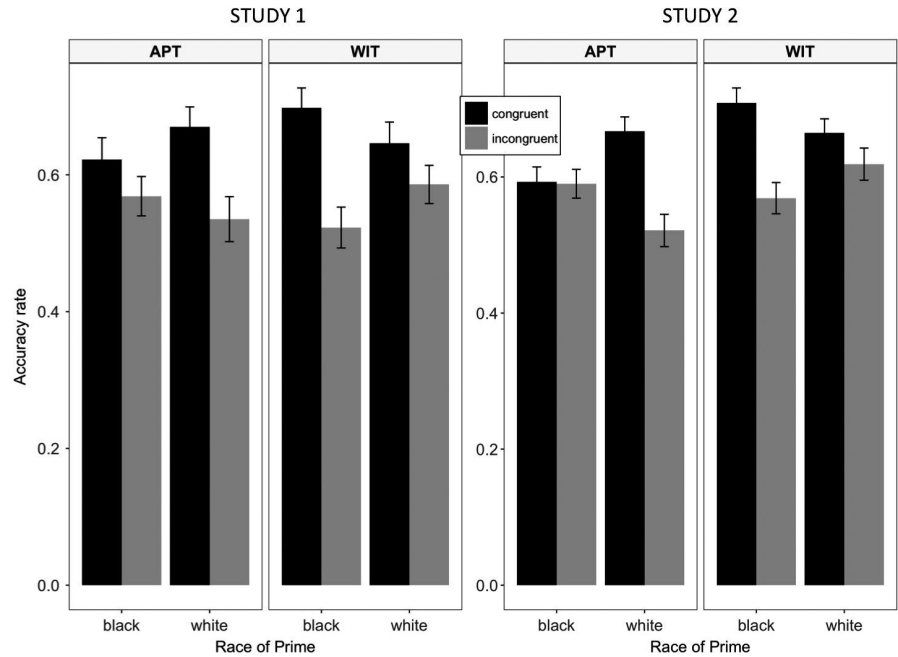
Abbreviations: APT, Affective Priming Task; WIT, Weapons Identification Task.

that a typical pattern of racial bias was evident in both tasks in that different race primes differentially influenced accuracy in categorizing the subsequent targets.

3.1.2 | Comparison across tasks

Response behavior across the tasks was compared in three ways. First, error rates from both tasks were subjected to a 2 (Prime: Black face, White face) \times 2 (Target: bias-congruent, bias-incongruent) \times 2 (Task: WIT, APT) rANOVA. A significant Prime \times Target \times Task interaction emerged, $F(1, 96) = 26.86, p < .001, \eta^2_p = 0.22$, (90% CI: [0.11, 0.33]), indicating that patterns of race bias differed across the tasks. As shown in Figure 1, the

FIGURE 1 Accuracy rate as a function of task and trial type. Congruent targets following Black faces are negative words in the APT and guns in the WIT. Congruent targets following White faces are positive words in the APT and tools in the WIT. Error bars represent standard error



difference in accuracy in the WIT between Black-congruent (gun) and Black-incongruent (tool) trials ($\eta^2_p = 0.51$) was larger than the difference between White-congruent (tool) and White-incongruent (gun) trials ($\eta^2_p = 0.08$).⁵ This pattern was reversed in the APT: the difference between White-congruent (positive) and White-incongruent (negative) trials ($\eta^2_p = 0.30$) was larger than the difference between Black-congruent (negative) and Black-incongruent (positive) trials ($\eta^2_p = 0.05$).

Second, response accuracy bias scores were created for each individual for each task as $p(\text{errors}|\text{incongruent}) - p(\text{errors}|\text{congruent})$, consistent with previous research (Ito et al., 2015; Payne, 2005; Wittenbrink, Judd, & Park, 1997), and were mean-centered and z-transformed. Response accuracy bias scores for the WIT were regressed onto response accuracy bias scores for the APT, revealing a modest association, $\beta = 0.23$, $p = .022$, $R^2 = .05$.

Third, PDP-Auto and PDP-Control estimates were compared across tasks. Standardized task-wide PDP-Control estimates from the APT were regressed onto standardized PDP-Control estimates from the WIT, revealing a strong association, $\beta = 0.60$, $p < .001$. Similar analyses using the PDP-Auto estimates showed what seemed to be a much weaker association across tasks, $\beta = 0.09$, $p = .361$. To compare the relative magnitude of PDP-Auto and PDP-Control associations, we used a mixed model in which PDP estimates from the WIT predicted the estimates from the APT, with type of estimate (*Auto* or *Control*) included as a predictor. The intercept was allowed to vary by participant to account for repeated observations. This model produced a significant Estimate \times Task

interaction, $\beta = 0.51$, $p < .001$, $f^2 = 0.10$,⁶ indicating that the association between PDP-Control estimates across the two tasks was significantly larger than the association between PDP-Auto estimates (Figure 2). A post-hoc power simulation estimated 98% power to detect an interaction of this size, given the sample size.⁷

3.2 | Study 2

3.2.1 | Preliminary analyses

Mean error rates (and SDs) for each condition of the WIT and APT are given in Table 2 and displayed in Figure 1. A significant congruency effect (i.e., Prime \times Target interaction) was evident in the WIT error rates, $F(1, 203) = 118.3$, $p < .001$, $\eta^2_p = 0.37$ (90% CI: [0.28, 0.44]). Guns were categorized more accurately than tools following Black faces, $F(1, 203) = 119.5$, $p < .001$, $\eta^2_p = 0.37$ (90% CI: [0.29, 0.44]), whereas tools were categorized more accurately than guns following White faces, $F(1, 203) = 8.77$, $p = .003$, $\eta^2_p = 0.04$ (90% CI: [0.01, 0.09]). The congruency effect also was significant for APT error rates, $F(1, 200) = 53.7$, $p < .001$, $\eta^2_p = 0.21$ (90% CI: [0.13, 0.29]). Accuracy categorizing negative and positive words following Black faces did not differ, $F(1, 200) = 0.00$, $p = .954$, $\eta^2_p = 0.00$, but positive words were categorized more accurately than negative words following White faces, $F(1, 200) = 102.1$, $p < .001$, $\eta^2_p = 0.34$ (90% CI: [0.25, 0.41]). Thus, patterns of response bias in Study 2 replicated those in Study 1.

⁶As suggested in Lorah (2018), Cohen's f^2 was used as a measure of effect size for fixed effects estimated by a mixed model. Cohen's f^2 was calculated using conditional R^2 (Nakagawa & Schielzeth, 2013) using the MuMIn package in R. Satterthwaite approximations were used to estimate degrees of freedom and to obtain two-tailed p -values.

⁷The simr R package was used to conduct the power simulation for the mixed model (Green & MacLeod, 2016).

⁵Although arguably tools are not stereotypically associated with Whites, previous research often finds a facilitation of tool responses following White faces (e.g., Payne, 2005). Therefore, we consider White-tool trials to be bias-congruent. This issue is considered further in the discussion.

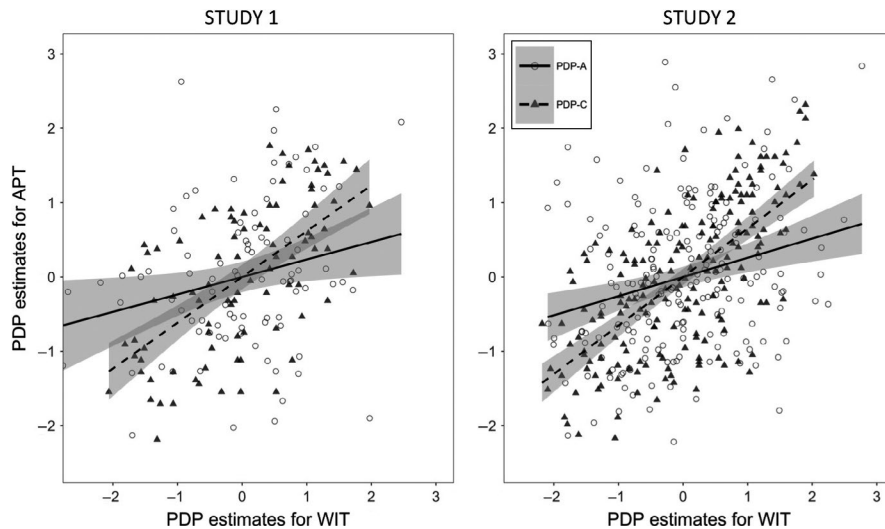


FIGURE 2 Slopes representing the correspondence between standardized task-wide automatic (PDP-Auto or PDP-A) and controlled (PDP-Control or PDP-C) processing estimates across tasks

3.2.2 | Comparison across tasks

Outcomes from the tasks were compared in the same three ways as in Study 1. When comparing patterns of error rates across the tasks, a significant Prime \times Target \times Task interaction emerged, $F(1, 200) = 97.32$, $p < .001$, $\eta^2_p = 0.33$ (90% CI: [0.24, 0.40]). As in Study 1, the interaction was characterized by different patterns of implicit bias across the tasks, such that the two-way interaction in the WIT was driven primarily by Black-prime trials, whereas the two-way interaction in the APT was driven primarily by White-prime trials. The correlation between response accuracy bias on the WIT and the APT was larger than in Study 1, $\beta = 0.40$, $p < .001$, $R^2 = .16$. Lastly, PDP-Auto and PDP-Control estimates were compared across the tasks, replicating patterns found in Study 1. PDP-Control estimates from the APT and the WIT were highly correlated, $\beta = 0.66$, $p < .001$, whereas PDP-Auto estimates seemed to be less highly correlated across tasks, $\beta = 0.30$, $p < .001$. As in Study 1, we used a mixed model in which PDP estimates from the WIT predicted PDP estimates from the APT, with type of estimate (*Auto* or *Control*) included as a predictor. The intercept was allowed to vary by participant to account for repeated observations. This model revealed a significant Estimate \times Task interaction, $\beta = 0.35$, $p < .001$, $f^2 = 0.02$, indicating a significantly larger association between PDP-Control estimates across the two tasks than the association between PDP-Auto estimates (Figure 2). A post-hoc power simulation estimated 98% power to detect an interaction of this size, given the sample size.

3.3 | Corroboration of results with the ABC model

One major limitation of the PDP is that it provides a relatively simplistic view of the processes involved in implicit bias task performance and does not account for response tendencies or guessing. Instead, the PDP assumes that when neither automatic nor controlled processes are active, the tool response is chosen in the Black-prime trials and the gun response is chosen in the White-prime trials (see Payne & Bishara, 2009; Sherman, 2008). Thus, the parameter

theorized to represent purely automatic processes is confounded by response and/or guessing bias (e.g., a tendency to use the response key of the dominant hand or a tendency to use the gun response). To account for response tendencies, the ABC model has been proposed as an alternative to the PDP and estimates a parameter representing response bias or guessing (B) that is constrained to be equal across different prime types (Stahl & Degner, 2007). To ensure our pattern of results was not an artifact of the confound between response bias and automatic processes, we additionally fit the ABC model to Study 1 and Study 2 data. We fit the ABC model to data in both tasks using the TreeBUGS R package (Heck, Arnold, & Arnold, 2018), which employs a Bayesian estimation approach to fit a multilevel extension of the model that treats subjects and items as random factors for each model parameter (Klauer, 2010). In the model for each task, we estimated four parameters: a separate A parameter for each type of prime, which theoretically represents the automatic activation of associations relevant to that prime, a C parameter, which represents controlled processing of the target, and a B parameter, which represents guessing or response bias.

In both studies, correspondence in the A and C parameters estimated by the ABC model across tasks showed a similar pattern of results as the PDP. The correspondence in ABC-C estimates across tasks in both Study 1, $\rho = 0.63$ (95% Bayesian Confidence Interval [BCI]: [0.47, 0.76]), and Study 2, $\rho = 0.69$ (95% BCI: [0.59, 0.77]), was similar to the correspondence in PDP-Control estimates across tasks (β s = 0.60 and 0.66 in Study 1 and Study 2, respectively). Correspondence in ABC-A estimates associated with Black primes across tasks in Study 1, $\rho = 0.23$ (95% BCI: [-0.54, 0.76]), and Study 2, $\rho = 0.53$ (95% BCI: [-0.31, 0.82]), and the correspondence in ABC-A estimates associated with White primes across tasks in Study 1, $\rho = 0.25$ (95% BCI: [-0.36, 0.68]), and Study 2, $\rho = 0.48$ (95% BCI: [-0.08, 0.78]), was smaller than correspondence between ABC-C estimates in each task, similar to the pattern seen with PDP-Control and PDP-Auto estimates. The BCIs are quite large for the estimates of correspondence between ABC-A estimates across tasks and overlap with the estimates of correspondence between ABC-C

estimates across tasks, so we cannot conclude that the ABC-A estimates correspond to a significantly smaller degree than the ABC-C estimates; however, estimates of the ABC-C correspondence across tasks are significantly larger than 0 whereas estimates of the ABC-A correspondence across tasks are not significantly different from 0. Importantly, patterns in correspondence of ABC estimates across tasks are consistent with patterns seen with PDP estimates.

4 | DISCUSSION

The primary purpose of these studies was to test the extent to which estimates of automatic and control-related processes correspond across two sequential priming tasks with identical structural features but tapping different kinds of implicit associations with race. Patterns of response accuracy bias and PDP estimates of automaticity and control were largely similar across the two experiments, increasing our confidence in the replicability of these patterns, although some potentially important differences in the magnitude of associations also emerged across the two studies.

Despite structural similarity between the two tasks, patterns of bias revealed by the WIT (stereotypic bias) and APT (evaluative bias) were markedly different. Whereas bias in the WIT was driven by larger differences between Black-congruent and Black-incongruent trials, bias in the APT was driven by larger differences between White-congruent and White-incongruent trials, suggesting the WIT appears to access outgroup stereotypes whereas the APT appears to access ingroup preferences. A pattern indicating Black prime-driven bias in the WIT is consistent with a response-mapping account (Scherer & Lambert, 2009), which posits that extreme stimuli that are strongly associated with a particular response displace the rating of less extreme stimuli onto the other available response. In the WIT, the strength of the stereotypical association between Black men and guns likely displaces the “tool” response to the White primes, producing an apparent association between White primes and tools ungrounded in cultural stereotypes.

In contrast, differential responding to the targets was more evident following White primes than Black primes in the APT. Quad model estimates from the IAT show a similar pattern, such that estimates of White-positive association activation are larger than estimates of Black-negative associations (e.g., Beer et al., 2008; Calanchini, Gonsalkorale, Sherman, & Klauer, 2013; Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Gonsalkorale, Allen, Sherman, & Klauer, 2010; Gonsalkorale, Sherman, Allen, Klauer, & Amodio, 2011; Gonsalkorale, Sherman, & Klauer, 2009). The current studies extend this base of evidence by demonstrating a similar effect using a sequential priming task, showing a stronger association between Whites and positive constructs than between Blacks and negative constructs in majority-White, emerging adult samples.

Of greater interest here, comparison of PDP estimates revealed that estimates of automatic associations derived from the two tasks were only modestly correlated in both studies, consistent with the idea that those estimates reflect the influence of different types of

automatic associations stemming from different learning experiences and rooted in distinct neurophysiological systems (Amodio, 2019). One implication of this finding, derived from the Multiple Memory Systems model of social cognition (Amodio, 2019; Amodio & Ratner, 2011), is that individual differences in the influence of evaluative bias might indicate very little concerning variability in semantic or stereotypic bias, and vice versa (see Osborne & Sibley, 2017, for an alternative approach to identifying individual differences in types of bias).

However, despite differences in their underlying nature, in some cases semantic and evaluative associations may not be completely decoupled. To many people, for example, handguns are not merely semantically associated with young Black men but also have a negative connotation (although not everyone; see Bartholow, Anderson, Carnagey, & Benjamin, 2005), which confounds the gun/tool stereotypic association with negative evaluations (Judd, Blair, & Chapleau, 2004). Thus, the degree to which an individual has both negative evaluations of a group and negative stereotypic associations with that group could result in a higher correspondence in estimates of automatic processing from separate tasks. In contrast, stimuli with less conceptual overlap between stereotypic and evaluative associations (for example, comparing negative evaluative associations with positive stereotypic associations between African-Americans and music or athleticism) may result in lower correspondence of automatic estimates (see Calanchini et al., 2014).

At the same time, it is important to acknowledge that estimates of automatic processing can show low correlations across tasks meant to tap the same form of bias. For example, Ito et al. (2015) administered three tasks tapping the stereotypic association between Blacks and gun violence, including two sequential priming tasks with very similar stimuli and timing parameters, and reported correlations of $r \leq .16$ in their PDP-Auto estimates. Thus, even tasks meant to assess the same form of bias produce estimates that are more divergent than convergent, suggesting that even latent variable approaches meant to overcome the limitations of single-indicator approaches (e.g., Cunningham, Preacher, & Banaji, 2001; Ito et al., 2015; Klauer, Schmitz, Teige-Mocigemba, & Voss, 2010) can provide only limited understanding of how learned associations affect biased responses across different tasks.

In contrast, the correlation between the PDP-Control estimates derived from the two tasks was relatively large ($r_s = .60$ and $.66$ in Studies 1 and 2, respectively), owing to the similarity in the tasks' structural features and response requirements and despite the differing associations on which their responses depended. These correlations are similar in magnitude to those reported in Ito et al. (2015) for the WIT and FPST ($r = .61$), despite the fact that the tasks used here were designed to tap different forms of bias, and similar to estimates of correspondence in controlled processing from the ABC model ($\rho_s = 0.63$ and 0.69). Thus, the current findings extend those of Ito et al. by showing that estimates of controlled processing across sequential priming tasks are relatively stable regardless of the specific types of automatic associations they measure. The high correspondence in PDP-Control estimates between tasks in the current studies suggests the same type of control is indexed by PDP-Control

in these tasks (suggested to be Common EF in Ito et al., 2015), and contributes to gathering evidence that task parameters and structure are important in determining the type of control-related processes that are engaged. A similar pattern would be expected for two IATs that separately measure stereotypic and evaluative associations—given a similar task structure, control-related processes are expected to be similar across IATs, whereas automatic associations are expected to differ (see Amodio & Devine, 2006; Calanchini et al. 2014).

An important additional aim of this research was to attempt to replicate patterns of association across two commonly used implicit bias tasks in two studies using very similar procedures but differing samples. Although observing both stereotypic and evaluative bias in sequential priming tasks is commonplace (see Cameron, Brown-Iannuzzi, & Payne, 2012), and although some previous studies have shown that measures of bias or estimates of automatic processing vary according to conceptual overlap within versions of the IAT (Amodio & Devine, 2006; Calanchini et al., 2014), no previous studies have specifically examined the extent to which estimates of bias and automatic associations correspond across sequential priming tasks measuring stereotypic and evaluative constructs. Thus, and given recent calls for replication of findings in social cognition (see Brandt et al., 2014; Spellman, 2013), we felt it was important to replicate findings across two very similar studies. The overall conclusions indicated by the data were quite similar in the two studies, namely, (a) that stereotypic accuracy bias was driven mainly by anti-Black stereotypes, whereas evaluative accuracy bias was driven mainly by pro-White evaluations; and (b) that estimates of automatic processing were correlated much less strongly than estimates of controlled processing. However, the second study showed stronger associations across tasks in both accuracy bias and automatic processing estimates than did the first study (though only the latter of these associations was significantly larger in Study 2, see Appendix S1). Given that the sample was twice as large in the second study compared to the first, this difference could be attributable to reduced error variance in the Study 2 data (also see Stanley & Spence, 2014). Regardless of the specific reason for the relatively small differences across the two studies' findings, we agree with the recommendations of Stanley and Spence, who advocate moving away from a focus on verification or refutation of previous results with newer findings and toward a cumulative, meta-analytic approach to evidence.

In conclusion, evidence has been mounting that different implicit bias measures tap different underlying automatic associations (Amodio, 2014; Amodio & Devine, 2006) and engage different types of control-related processes (Ito et al., 2015; Klauer et al., 2010; Sherman, 2008). The current findings add to this literature and have a number of implications for researchers' decisions concerning measurement of implicit bias. Perhaps most importantly, the current findings support and extend recent work indicating that a researcher's choice of bias tasks should be driven by theoretical considerations of the specific type of bias that is most relevant to measure, as well as the type(s) and degree of control-related processing likely to influence that measurement. Ultimately, many researchers are

interested in implicit bias to the extent that it can explain or predict real-world behavior. Despite evidence that stereotypic and evaluative bias are separate constructs and contribute to different types of behavior (e.g., Amodio, 2019; Amodio & Devine, 2006), researchers often rely upon a single implicit bias measure to examine real-life consequences of bias or to test the effectiveness of a bias-reduction training program (e.g., Cameron, Brown-Iannuzzi, & Payne, 2012; Devine, Forscher, Austin, & Cox, 2012; Greenwald, Poehlman, Uhlmann, & Banaji, 2009; Hausmann & Ryan, 2004; Hehman et al., 2017; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Lai et al., 2016; Maister, Sebanz, Knoblich, & Tsakiris, 2013; Peck, Seinfeld, Aglioti, & Slater, 2013). Use of a single implicit bias measure, particularly if selected for convenience or because of researchers' own preferences, provides a simplistic picture of the relationships between race-related attitudes and stereotypes and other relevant constructs.

Additionally, the specific structure of implicit bias tasks must be considered carefully. The use of similar timing parameters and imposition of a response deadline in the current studies produced similarity across tasks in the influence of control-related processes, despite differences in the automatic associations they measured. Researchers should consider administering several different bias tasks and using a latent variable approach to eliminate task-specific variance, particularly when implicit bias is used as a predictor or outcome of real-world behaviors or interventions (Cunningham et al., 2001; Ito et al., 2015; Klauer et al., 2010). Careful consideration of the automatic and controlled processes that influence behavior in implicit bias tasks are beneficial in the study of the real-life consequences of implicit bias and interventions to reduce implicit bias.

FUNDING INFORMATION

This research was supported in part by a Life Sciences Graduate Fellowship from the University of Missouri. Preparation of this manuscript was supported by the National Institute on Alcohol Abuse and Alcoholism [grant R01 AA020970] and the National Science Foundation [grant number 1529788].

CONFLICT OF INTEREST

The authors have no conflicts of interest.

TRANSPARENCY STATEMENT

Data and analyses can be found at <https://github.com/hiv8r3/ComparingBias-project>.

ORCID

Hannah I. Volpert-Esmond  <https://orcid.org/0000-0002-9181-9710>

REFERENCES

- Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience*, 15(10), 670–682. <https://doi.org/10.1038/nrn3800>

- Amodio, D. M. (2019). Social Cognition 2.0: An interactive memory systems account. *Trends in Cognitive Sciences*, 23(1), 21–33. <https://doi.org/10.1016/j.tics.2018.10.002>
- 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(4), 652–661. <https://doi.org/10.1037/0022-3514.91.4.652>
- Amodio, D. M., & Hamilton, H. K. (2012). Intergroup anxiety effects on implicit racial evaluation and stereotyping. *Emotion*, 12(6), 1273–1280. <https://doi.org/10.1037/a0029016>
- Amodio, D. M., & Ratner, K. G. (2011). A memory systems model of implicit social cognition. *Current Directions in Psychological Science*, 20(3), 143–148. <https://doi.org/10.1177/0963721411408562>
- Banisek, R., Seise, J., & Zerbates, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für experimentelle Psychologie*, 48(2), 145–160. <https://doi.org/10.1026//0949-3946.48.2.145>
- Bartholow, B. D., Anderson, C. A., Carnagey, N. L., & Benjamin, A. J. (2005). Interactive effects of life experience and situational cues on aggression: The weapons priming effect in hunters and nonhunters. *Journal of Experimental Social Psychology*, 41(1), 48–60. <https://doi.org/10.1016/j.jesp.2004.05.005>
- Bartholow, B. D., Riordan, M. A., Saults, J. S., & Lust, S. A. (2009). Psychophysiological evidence of response conflict and strategic control of responses in affective priming. *Journal of Experimental Social Psychology*, 45(4), 655–666. <https://doi.org/10.1016/j.jesp.2009.02.015>
- 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(4), 775–783. <https://doi.org/10.1016/j.neuroimage.2008.08.033>
- Brandt, M. J., Ilzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., ... van't Veer, A. (2014). The replication recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, 50, 217–224. <https://doi.org/10.1016/j.jesp.2013.10.005>
- Brendl, C. M., Markman, A. B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology*, 81(5), 760–773.
- Brownstein, M., Madva, A., & Gawronski, B. (2019). What do implicit measures measure? *Wiley Interdisciplinary Reviews: Cognitive Science*, 10, e1501. <https://doi.org/10.1002/wcs.1501>
- Calanchini, J., Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2013). Counter-prejudicial training reduces activation of biased associations and enhances response monitoring. *European Journal of Social Psychology*, 43(5), 321–325. <https://doi.org/10.1002/ejsp.1941>
- Calanchini, J., Sherman, J. W., Klauer, K. C., & Lai, C. K. (2014). Attitudinal and non-attitudinal components of IAT performance. *Personality and Social Psychology Bulletin*, 40(10), 1285–1296. <https://doi.org/10.1177/0146167214540723>
- Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: A meta-analysis of associations with behavior and explicit attitudes. *Personality and Social Psychology Review*, 16(4), 330–350. <https://doi.org/10.1177/1088868312440047>
- Chaiken, S. & Trope, Y. (Eds.) (1999). *Dual process theories in social psychology*. New York, NY: Guilford Press.
- Cocchetti, D. V. (1999). Sample size requirements for increasing the precision of reliability estimates: Problems and proposed solutions. *Journal of Clinical and Experimental Neuropsychology*, 21(4), 567–570. <https://doi.org/10.1076/jcen.21.4.567.886>
- Conroy, 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(4), 469–487.
- 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(6), 1314–1329. <https://doi.org/10.1037/0022-3514.83.6.1314>
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science*, 12(2), 163–170. <https://doi.org/10.1111/1467-9280.00328>
- De Houwer, J. (2001). A structural and process analysis of the implicit association test. *Journal of Experimental Social Psychology*, 37(6), 443–451. <https://doi.org/10.1006/jesp.2000.1464>
- De Houwer, J. (2019). Implicit bias is behavior: A functional-cognitive perspective on implicit bias. *Perspectives on Psychological Science*, 14(5), 835–840. <https://doi.org/10.1177/1745691619855638>
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, 135(3), 347–368. <https://doi.org/10.1037/a0014211>
- Degner, J. (2009). On the (un-)controllability of affective priming: Strategic manipulation is feasible but can possibly be prevented. *Cognition and Emotion*, 23(2), 327–354. <https://doi.org/10.1080/02699930801993924>
- Degner, J., & Wentura, D. (2010). Automatic prejudice in childhood and early adolescence. *Journal of Personality and Social Psychology*, 98(3), 356–374. <https://doi.org/10.1037/a0017993>
- Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. L. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, 48(6), 1267–1278. <https://doi.org/10.1016/j.jesp.2012.06.003>
- Dovidio, J. F., Gaertner, S. L., & Pearson, A. R. (2017). Aversive racism and contemporary bias. In C. G. Sibley & F. K. Barlow (Eds.), *The Cambridge handbook of the psychology of prejudice* (pp. 267–294). New York, NY: Cambridge University Press.
- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82(1), 62–68. <https://doi.org/10.1037/0022-3514.82.1.62>
- 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(6), 1013–1027. <https://doi.org/10.1037/0022-3514.69.6.1013>
- Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, 86, 186–204. <https://doi.org/10.1016/j.cortex.2016.04.023>
- Gawronski, B., & Hahn, A. (2019). Implicit measures: Procedures, use, and interpretation. In H. Blanton & G. D. Webster (Eds.), *Foundations of social psychological measurement*. New York, NY: Taylor & Francis.
- Gonsalkorale, K., Allen, T. J., Sherman, J. W., & Klauer, K. C. (2010). Mechanisms of group membership and exemplar exposure effects on implicit attitudes. *Social Psychology*, 41(3), 158–168. <https://doi.org/10.1027/1864-9335/a000023>
- 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(11), 1534–1545. <https://doi.org/10.1177/0146167211414064>
- Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology*, 45(2), 410–414. <https://doi.org/10.1016/j.jesp.2008.11.004>
- Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>

- Greenwald, A. G., & Banaji, M. R. (2017). The implicit revolution: Reconciling the relation between conscious and unconscious. *American Psychologist*, 72(9), 861–871. <https://doi.org/10.1037/amp0000238>
- Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of Personality and Social Psychology*, 108(4), 553–561. <https://doi.org/10.1037/pspa0000016>
- 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(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97(1), 17–41. <https://doi.org/10.1037/a0015575>
- Hall, W. J., Chapman, M. V., Lee, K. M., Merino, Y. M., Thomas, T. W., Payne, B. K., ... Coyne-Beasley, T. (2015). Implicit racial/ethnic bias among health care professionals and its influence on health care outcomes: A systematic review. *American Journal of Public Health*, 105(12), e60–e76. <https://doi.org/10.2105/AJPH.2015.302903>
- Han, H. A., Olson, M. A., & Fazio, R. H. (2006). The influence of experimentally created extrapersonal associations on the Implicit Association Test. *Journal of Experimental Social Psychology*, 42(3), 259–272. <https://doi.org/10.1016/j.jesp.2005.04.006>
- Hausmann, L. R. M., & Ryan, C. S. (2004). Effects of external and internal motivation to control prejudice on implicit prejudice: The mediating role of efforts to control prejudiced responses. *Basic and Applied Social Psychology*, 26(2–3), 215–225. <https://doi.org/10.1080/01973533.2004.9646406>
- Heck, D. W., Arnold, N. R., & Arnold, D. (2018). TreeBUGS: An R package for hierarchical multinomial-processing-tree modeling. *Behavior Research Methods*, 50(1), 264–284. <https://doi.org/10.3758/s13428-017-0869-7>
- Helman, E., Flake, J. K., & Calanchini, J. (2017). Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *Social Psychological and Personality Science*, 9(4), 393–401. <https://doi.org/10.1177/1948550617711229>
- 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(10), 1369–1385. <https://doi.org/10.1177/0146167205275613>
- Hütter, M., & Klauer, K. C. (2016). Applying processing trees in social psychology. *European Review of Social Psychology*, 27(1), 116–159. <https://doi.org/10.1080/10463283.2016.1212966>
- Ito, T. A., Friedman, N. P., Bartholow, B. D., Correll, J., Loersch, C., Altamirano, L. J., & Miyake, A. (2015). Toward a comprehensive understanding of executive cognitive function in implicit racial bias. *Journal of Personality and Social Psychology*, 108(2), 187–218. <https://doi.org/10.1037/a0038557>
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30(5), 513–541. [https://doi.org/10.1016/0749-596X\(91\)90025-F](https://doi.org/10.1016/0749-596X(91)90025-F)
- Judd, C. M., Blair, I. V., & Chapleau, K. M. (2004). Automatic stereotypes vs. automatic prejudice: Sorting out the possibilities in the Payne (2001) weapon paradigm. *Journal of Experimental Social Psychology*, 40(1), 75–81. [https://doi.org/10.1016/S0022-1031\(03\)00063-5](https://doi.org/10.1016/S0022-1031(03)00063-5)
- Kawakami, K., & Dovidio, J. F. (2001). The reliability of implicit stereotyping. *Personality and Social Psychology Bulletin*, 27(2), 212–225. <https://doi.org/10.1177/0146167201272007>
- Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika*, 75(1), 70–98. <https://doi.org/10.1007/s11336-009-9141-0>
- 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 implicit association test effect. *Personality and Social Psychology Bulletin*, 31(2), 208–217. <https://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(3), 595–619. <https://doi.org/10.1080/17470210903076826>
- Klinger, M. R., Burton, P. C., & Pitts, G. S. (2000). Mechanisms of unconscious priming: I. Response competition, not spreading activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(2), 441–455. <https://doi.org/10.1037/0278-7393.26.2.441>
- Kurdi, B., Seitchik, A. E., Axt, J. R., Carroll, T. J., Karapetyan, A., Kaushik, N., ... Banaji, M. R. (2019). Relationship between the Implicit Association Test and intergroup behavior: A meta-analysis. *American Psychologist*, 74(5), 569–586. <https://doi.org/10.1037/amp0000364>
- LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: A mixed-trial fMRI study. *Neuron*, 20(5), 937–945. [https://doi.org/10.1016/S0896-6273\(00\)80475-4](https://doi.org/10.1016/S0896-6273(00)80475-4)
- Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., ... Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, 145(8), 1001–1016. <https://doi.org/10.1037/xge0000179>
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, 23(1), 155–184. <https://doi.org/10.1146/annurev.neuro.23.1.155>
- Livingston, R. W., & Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology*, 82(1), 5–18. <https://doi.org/10.1037/0022-3514.82.1.5>
- Lorah, J. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education*, 6(1), 6–8. <https://doi.org/10.1186/s40536-018-0061-2>
- Machery, E. (2017). Do indirect measures of biases measure traits or situations? *Psychological Inquiry*, 28(4), 288–291. <https://doi.org/10.1080/1047840X.2017.1373557>
- Maister, L., Sebanz, N., Knoblich, G., & Tsakiris, M. (2013). Experiencing ownership over a dark-skinned body reduces implicit racial bias. *Cognition*, 128(2), 170–178. <https://doi.org/10.1016/j.cognition.2013.04.002>
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58(1), 25–45. <https://doi.org/10.1146/annurev.psych.57.102904.190143>
- McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, 114(2), 159–197. <https://doi.org/10.1037//0096-3445.114.2.159>
- 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(1), 8–14. <https://doi.org/10.1177/0963721411429458>
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>
- Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extrapersonal associations on the implicit association test: Personalizing the IAT. *Journal of Personality and Social Psychology*, 86(5), 653–667. <https://doi.org/10.1037/0022-3514.86.5.653>
- Osborne, D., & Sibley, C. G. (2017). Identifying “types” of ideologies and intergroup biases: Advancing a person-centred approach to social

- psychology. *European Review of Social Psychology*, 28(1), 288–332. <https://doi.org/10.1080/10463283.2017.1379265>
- Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, 105(2), 171–192. <https://doi.org/10.1037/a0032734>
- 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(2), 181–192. <https://doi.org/10.1037/0022-3514.81.2.181>
- Payne, B. K. (2005). Conceptualizing control in social cognition: How executive functioning modulates the expression of automatic stereotyping. *Journal of Personality and Social Psychology*, 89(4), 488–503. <https://doi.org/10.1037/0022-3514.89.4.488>
- Payne, B. K., & Bishara, A. J. (2009). An integrative review of process dissociation and related models in social cognition. *European Review of Social Psychology*, 20(1), 272–314. <https://doi.org/10.1080/10463280903162177>
- 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(4), 384–396. [https://doi.org/10.1016/S0022-1031\(02\)00006-9](https://doi.org/10.1016/S0022-1031(02)00006-9)
- Payne, B. K., Vuletic, H. A., & Lundberg, K. B. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, 28(4), 233–248. <https://doi.org/10.1080/1047840X.2017.1335568>
- Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, 22(3), 779–787. <https://doi.org/10.1016/j.concog.2013.04.016>
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, 75(3), 811–832. <https://doi.org/10.1037/0022-3514.75.3.811>
- Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42–55. <https://doi.org/10.1038/nrn.2016.150>
- Scherer, L. D., & Lambert, A. J. (2009). Contrast effects in priming paradigms: Implications for theory and research on implicit attitudes. *Journal of Personality and Social Psychology*, 97(3), 383–403. <https://doi.org/10.1037/a0015844>
- Sherman, J. W. (2008). Controlled influences on implicit measures: Confronting the myth of process-purity and taming the cognitive monster. In R. E. Petty, R. H. Fazio & P. Brinol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 391–426). New York, NY: Psychology Press.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119(1), 3–22.
- Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, 4(2), 108–131. https://doi.org/10.1207/S15327957PSPR0402_01
- Spellman, B. A. (2013). There is no such thing as replication, but we should do it anyway. *European Journal of Personality*, 27(2), 136–137.
- Stahl, C., & Degner, J. (2007). Assessing automatic activation of valence. *Experimental Psychology*, 54(2), 99–112. <https://doi.org/10.1027/1618-3169.54.2.99>
- Stanley, D. J., & Spence, J. R. (2014). Expectations for replications: Are yours realistic? *Perspectives on Psychological Science*, 9(3), 305–318. <https://doi.org/10.1177/1745691614528518>
- 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(2), 164–168. <https://doi.org/10.1111/j.1467-9280.2009.02274.x>
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8(3), 220–247. https://doi.org/10.1207/s15327957pspr0803_1
- Teige-Mocigemba, S., Klauer, K. C., & Sherman, J. W. (2010). A practical guide to implicit association tests and related tasks. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition* (pp. 117–139). New York, NY: Guilford Press.
- Wentura, D., & Degner, J. (2010). A practical guide to sequential priming and related tasks. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 95–115). New York, NY: Guilford Press.
- Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, 72(2), 262–274. <https://doi.org/10.1037/0022-3514.72.2.262>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Volpert-Esmond HI, Scherer LD, Bartholow BD. Dissociating automatic associations: Comparing two implicit measurements of race bias. *Eur J Soc Psychol*. 2019;00:1–13. <https://doi.org/10.1002/ejsp.2655>