Predicting Ethnic and Racial Discrimination: A Meta-Analysis of IAT Criterion Studies

Frederick L. Oswald
Rice University

Gregory Mitchell
University of Virginia

Hart Blanton
University of Connecticut

James Jaccard
New York University

Philip E. Tetlock
University of Pennsylvania

This article reports a meta-analysis of studies examining the predictive validity of the Implicit Association Test (IAT) and explicit measures of bias for a wide range of criterion measures of discrimination. The meta-analysis estimates the heterogeneity of effects within and across 2 domains of intergroup bias (interracial and interethnic), 6 criterion categories (interpersonal behavior, person perception, policy preference, microbehavior, response time, and brain activity), 2 versions of the IAT (stereotype and attitude IATs), 3 strategies for measuring explicit bias (feeling thermometers, multi-item explicit measures such as the Modern Racism Scale, and ad hoc measures of intergroup attitudes and stereotypes), and 4 criterion-scoring methods (computed majority–minority difference scores, relative majority–minority ratings, minority-only ratings, and majority-only ratings). IATs were poor predictors of every criterion category other than brain activity, and the IATs performed no better than simple explicit measures. These results have important implications for the construct validity of IATs, for competing theories of prejudice and attitude–behavior relations, and for measuring and modeling prejudice and discrimination.

Keywords: Implicit Association Test, explicit measures of bias, predictive validity, discrimination, prejudice

Supplemental materials: http://dx.doi.org/10.1037/a0032734.supp

Although only 14 years old, the Implicit Association Test (IAT) has already had a remarkable impact inside and outside academic psychology. The research article introducing the IAT (Greenwald, McGhee, & Schwartz, 1998) has been cited over 2,600 times in PsycINFO and over 4,300 times in Google Scholar, and the IAT is now the most commonly used implicit measure in psychology. Trade book translators of psychological research cite IAT findings as evidence that human behavior is much more under the control of unconscious forces—and much less under control of volitional forces—than lay intuitions would suggest (e.g., Malcolm Gladwell's 2005 bestseller, *Blink*; Shankar Vedantam's 2010 *The Hidden Brain*; and Banaji and Greenwald's 2013 *Blindspot*). Observers of the political scene invoke IAT-based research conclusions about implicit bias as explanations for a wide range of controversies, from vote counts in presidential primaries (Parks & Rachlinski, 2010) to racist outbursts by celebrities (Shermer, 2006) to outrage over a *New Yorker* magazine cover depicting Barack Obama as a Muslim (Banaji, 2008). In courtrooms, expert witnesses invoke IAT research to support the proposition that unconscious bias is a pervasive cause of employment discrimination (Greenwald, 2006; Scheck, 2004). Law professors (e.g., Kang, 2005; Page & Pitts, 2009; Shin, 2010) and sitting federal judges (Bennett, 2010) cite IAT research conclusions as grounds for changing laws. Indeed, the National Center for State Courts and the American Bar Association have launched programs to educate judges, lawyers, and court administrators on the dangers of im-

This article was published Online First June 17, 2013.

Fred L. Oswald, Department of Psychology, Rice University; Gregory Mitchell, School of Law, University of Virginia; Hart Blanton, Department of Psychology, University of Connecticut; James Jaccard, Center for Latino Adolescent and Family Health, Silver School of Social Work, New York University; Philip E. Tetlock, Wharton School of Business, University of Pennsylvania.

Fred L. Oswald, Gregory Mitchell, and Philip E. Tetlock are consultants for LASSC, LLC, which provides services related to legal applications of social science research, including research on prejudice and stereotypes. We thank Carter Lennon for her comments on an earlier version of the article and Dana Carney, Jack Glaser, Eric Knowles, and Laurie Rudman for their helpful input on data coding.

Correspondence concerning this article should be addressed to Frederick L. Oswald, Department of Psychology, Rice University, 6100 Main Street MS25, Houston, TX 77005-1827; to Gregory Mitchell, School of Law, University of Virginia, 580 Massie Road, Charlottesville, VA 22903-1738; or to Hart Blanton, Department of Psychology, University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT 06269-1020. E-mail: foswald@rice.edu or greg_mitchell@virginia.edu or hart.blanton@uconn.edu

 0022-3514/13/$12.00 DOI: 10.1037/a0032734

plicit bias in the legal system, and many of the lessons in these programs are drawn directly from the IAT literature (Drummond, 2011; Irwin & Real, 2010).

These applications of IAT research assume that the IAT predicts discrimination in real-world settings (Tetlock & Mitchell, 2009). Although only a handful of studies have examined the predictive validity of the IAT in field settings (e.g., Agerström & Rooth, 2011), many laboratory studies have examined the correlation between IAT scores and criterion measures of intergroup discrimination. The earliest IAT criterion studies were predicated on social cognitive theories that assign greater influence to implicit attitudes on spontaneous than deliberate responses to stimuli (e.g., Fazio, 1990; see Dovidio, Kawakami, Smoak, & Gaertner, 2009; Olson & Fazio, 2009). These investigations examined the correlation between IAT scores and the spontaneous, often subtle behaviors exhibited by majority-group members in interactions with minority-group members (e.g., facial expressions and body posture; cf. McConnell & Leibold, 2001; Richeson & Shelton, 2003; Vanman, Saltz, Nathan, & Warren, 2004). Other studies sought to go deeper, using such approaches as fMRI technology to identify the neurological origins of implicit biases and discrimination (e.g., Cunningham et al., 2004; Richeson et al., 2003). As the popularity of implicit bias as a putative explanation for societal inequalities grew (e.g., Blasi & Jost, 2006), criterion studies started examining the relation of IAT scores to more deliberate conduct, such as judgments of guilt in hypothetical trials, the treatment of hypothetical medical patients, and voting choices (e.g., Green et al., 2007; Greenwald, Smith, Sriram, Bar-Anan, & Nosek, 2009; Levinson, Cai, & Young, 2010).

In 2009, Greenwald, Poehlman, Uhlmann, and Banaji quantitatively synthesized 122 criterion studies across many domains in which IAT scores have been used to predict behavior, ranging from self-injury and drug use to consumer product preferences and interpersonal relations. They concluded that, "for socially sensitive topics, the predictive power of self-report measures was remarkably low and the incremental validity of IAT measures was relatively high" (Greenwald, Poehlman, et al., 2009, p. 32). In particular, "IAT measures had greater predictive validity than did self-report measures for criterion measures involving interracial behavior and other intergroup behavior" (Greenwald, Poehlman, et al., 2009, p. 28).

The Greenwald, Poehlman, et al. (2009) findings have potentially far-ranging theoretical, methodological, and even policy implications. First, these results appear to support the construct validity of the IAT. Because of controversies surrounding what exactly the IAT measures, a key test of the IAT's construct validity is whether it predicts relevant social behaviors (e.g., Arkes & Tetlock, 2004; Karpinski & Hilton, 2001; Rothermund & Wentura, 2004), and Greenwald, Poehlman, et al.'s findings suggest that this test has been passed. Second, Greenwald, Poehlman, et al.'s finding that the IAT predicted criteria across levels of controllability weighs against theories that assign implicit constructs greater influence on spontaneous than controlled behavior (e.g., Strack & Deutsch, 2004; see Perugini, Richetin, & Zogmaister, 2010). Third, the finding that the IAT outperformed explicit measures in socially sensitive domains, paired with the finding that both implicit and explicit measures showed incremental validity across domains, supports dual-construct theories of attitudes. It further argues in favor of the use of both implicit and explicit assessment, particularly when assessing attitudes or preferences involving sensitive topics. Finally, and most important, these findings appear to validate the concept of implicit prejudice as an explanation for social inequality and demonstrate that the IAT can be a useful predictor of who will engage in both subtle and not-so-subtle acts of discrimination against African Americans and other minorities. In short, Greenwald, Poehlman, et al. (2009) "confirms that implicit biases, particularly in the context of race, are meaningful" (Levinson, Young, & Rudman, 2012, p. 21). That confirmation in turn supports application of IAT research to the law and public policy, particularly with respect to the regulation of intergroup relations (see, e.g., Kang et al., 2012; Levinson & Smith, 2012).

The Need for a Closer Look at the Prediction of Intergroup Behavior

Although the findings reported by Greenwald, Poehlman, et al. (2009) have generated considerable enthusiasm, certain findings in their published report suggest that any conclusions about the satisfactory predictive validity of the IAT should be treated as provisional, especially when considered in light of findings reported in other relevant meta-analyses. First, Greenwald et al. found that the IAT did not outperform explicit measures for a number of sensitive topics (e.g., willingness to reveal drug use or true feelings toward intimate others), and explicit measures substantially outperformed IATs in the prediction of behavior and other criteria in several important domains. Indeed, in seven of the nine criterion domains examined by Greenwald et al. (gender/sex orientation preferences, consumer preferences, political preferences, personality traits, alcohol/drug use, psychological health, and close relationships), explicit measures showed higher correlations with criterion measures than did IAT scores, often by practically significant margins. Second, Greenwald et al.'s conclusion that the IAT and explicit measures appear to tap into different constructs and that explicit measures are less predictive for socially sensitive topics is at odds with meta-analytic findings by Hofmann, Gawronski, Gschwendner, Le, and Schmitt (2005) that implicit–explicit correlations were not influenced by social desirability pressures. Hofmann et al. concluded that IAT and explicit measures are systematically related and that variation in that relationship depends on method variance, the spontaneity of explicit measures, and the degree of conceptual correspondence between the measures.[1] Third, the low correlations between explicit measures of prejudice and criteria reported by Greenwald et al. (both rs = .12 for the race and other intergroup domains) are at odds with Kraus's (1995) estimate of the attitude–behavior correlation for explicit prejudice measures (r = .24) and a similar estimate by Talaska, Fiske, and Chaiken (2008; r = .26). These inconsistencies raise questions about the quality of the explicit measures of bias used in the IAT criterion studies. If explicit measures used in the IAT criterion studies had possessed the same predictive validity as measures considered by Kraus (1995) and Talaska et al. (2008), the IAT would not have outperformed the explicit measures in any domain. It is possible, however, given

[1] Cameron, Brown-Iannuzzi, and Payne (2012) noted that the use of different subjective coding methods may account for differences in meta-analytic results regarding social sensitivity as a moderator of the relation of implicit and explicit attitudes (see also Bar-Anan & Nosek, 2012).

the diverse ways that discrimination has been operationalized in the IAT criterion studies, that no explicit measures, regardless of how well constructed, could have achieved equivalent validity levels.

To better understand when and why the IAT and self-report measures differentially predict criteria, one must examine possible moderators of the construct–criterion relationship. Greenwald, Poehlman, et al. (2009) performed moderator analyses, but they focused on construct–criterion relations across criterion domains and did not report moderator results within criterion domains. Their cross-domain moderator results must be viewed cautiously for a number of reasons. First, as they note, "criterion domain variations were extensively confounded with several conceptual moderators" (Greenwald, Poehlman, et al., 2009, p. 24). Second, Greenwald, Poehlman, et al.'s meta-analytic method utilized a single effect size for each sample studied. As a result, studies using disparate criterion measures were assigned a single effect size, derived by averaging correlations across the criteria employed. Even if the criteria in a single study varied in terms of controllability or social desirability—and even if researchers sought to manipulate such factors across experimental conditions (e.g., Ziegert & Hanges, 2005)—every criterion in the study received the same score on the moderator of interest. Third, in the domains of interracial and other intergroup interactions, there was little variation across studies in the values assigned to key moderator variables (e.g., with one exception, the race IAT and explicit measures were given the same social desirability ratings whenever both types of measures were used in a study). Finally, inconsistencies were discovered in the moderator coding by Greenwald, Poehlman, et al., and it was therefore hard to understand and replicate some of their coding decisions (see online supplemental materials for details).[2]

The cumulative effect of these analytical and coding decisions was to obscure possible heterogeneity of effects connected to differences in the explicit measures used, the criterion measures used, and the methods used to score the criterion measures. In just the domain of interracial relations, criteria included such disparate indicators as the nonverbal treatment of a stranger, the endorsement of specific political candidates, and the results of fMRI scans recorded while respondents performed other laboratory tasks. These criteria were scored in a variety of ways that emphasize attitudes toward the majority group, the minority group, or both (i.e., absolute ratings of Black or White targets, ratings for White and Black targets on a common scale, or difference scores computed from separate ratings for White and Black targets). Substantive variability in performance on these criterion measures, as predicted by the IAT, different explicit measures, or differences in criterion scoring, were not open to scrutiny under the meta-analytic and moderator approaches adopted by Greenwald, Poehlman, et al. (2009).

Therefore, to address important theoretical and applied questions raised by the diverse findings from Greenwald, Poehlman, et al. (2009), and in particular to better understand the relation of implicit and explicit bias to discriminatory behavior, a new meta-analysis of the IAT criterion studies is needed. The existing meta-analytic literature on attitude–behavior relations does not answer these questions. The meta-analyses conducted by Kraus (1995) and Talaska et al. (2008) emphasized the relation of explicit measures of attitudes to prejudicial behavior. The meta-analysis by Cameron, Brown-Iannuzzi, and Payne (2012) examined the prediction of a wide range of behavior by explicit and implicit attitude measures, including prejudicial behavior, but it focused on sequential priming measures and did not examine the predictive validity of IATs.

A New Meta-Analysis of Ethnic and Racial Discrimination Criterion Studies

The present meta-analysis examines the predictive utility of the IAT in two of the criterion domains that were most strongly linked to the predictive validity of the IAT in Greenwald, Poehlman, et al. (2009)—Black–White relations and ethnic relations—and that understandably invoke strong applied interest (e.g., Kang, 2005; Levinson & Smith, 2012; Page & Pitts, 2009).[3] It provides a detailed comparison of the IAT and explicit measures of bias as predictors of different forms of discrimination within these two domains. It would be both scientifically and practically remarkable if the IAT and explicit measures of bias were equally good predictors of the many different criterion measures used as proxies for racial and ethnic discrimination in the studies, because the criterion measures cover a vast range of levels of analysis and employ very different assessment methods. By differentiating among the ways in which prejudice was operationalized within the criterion studies, we can examine heterogeneity of effects within and across categories, identify sources of heterogeneity, and answer a number of questions regarding the construct validity of the IAT and the nature of the relationship between behavior and intergroup bias measured implicitly and explicitly.

---

[2] Part of the difficulty lies, no doubt, in the inherent ambiguity that surrounds trying to place sometimes complex tasks and manipulations onto single dimensions of social-psychological significance after the fact. Consider, for example, the "degree of conscious control" associated with a criterion, one of the key moderators examined by Greenwald, Poehlman, et al. (2009). It is difficult to know how conscious control might differ for verbal versus nonverbal behaviors and how these responses might differ from self-reported social perceptions. It is similarly unclear how controllability of participant responses to a computer task might differ from the images taken of the brains of participants who are performing the same computer tasks. For a number of the moderators employed by Greenwald, Poehlman, et al., we were unable to produce scores that lined up with their scores or understand why our ratings differed.

[3] Greenwald, Poehlman, et al. (2009) concluded that the predictive validity of the IAT outperformed explicit measures in the White–Black race and other intergroup criterion domains. Our race domain parallels Greenwald, Poehlman, et al.'s White–Black race domain, with all studies examining bias against African Americans/Africans relative to White Americans/European Americans. Greenwald, Poehlman, et al.'s other intergroup domain included studies examining bias against ethnic groups, older persons, religious groups, and obese persons (i.e., Greenwald, Poehlman, et al.'s other intergroup domain appears to have been a catchall category rather than theoretically or practically unified). The wide range of groups placed under the other intergroup label by Greenwald, Poehlman, et al. risks combining bias phenomena that implicate very different social-psychological processes. Our analysis focuses on race discrimination and discrimination against ethnic groups and foreigners (i.e., national-origins discrimination), because ethnicity and race are characteristics that often involve observable differences that can be the basis of automatic categorization.

Moderators Examined and Questions Addressed

Criterion domain: Does the relationship between discrimination and scores on the IAT or explicit measures of bias vary as a function of the nature of the intergroup relation?

We differentiated between White–Black relations and interethnic relations in our coding of criterion studies to account for a possible source of variation in effects, but we did not have strong theoretical or empirical reasons to believe that criterion prediction would differ by the nature of the intergroup relation. IAT researchers often find score distributions that are interpreted as revealing high levels of bias against both African American and various ethnic minorities (e.g., Nosek et al., 2007; cf. Blanton & Jaccard, 2006), and these patterns were replicated within the criterion studies we examine. Reports of high levels of explicitly measured racial and ethnic bias are less common in the literature (e.g., Quillian, 2006; Sears, 2004a) and within the criterion studies we examine. Thus, we did not expect the pattern of construct–criterion relations to vary between the race and ethnicity domains.

Nature of the criterion: Does the relationship between discrimination and scores on the IAT or explicit measures of bias vary as a function of the manner in which discrimination is operationalized?

We placed criteria into one of six easily distinguishable categories of criterion measures used in the IAT criterion studies as indicators of discrimination: (a) *brain activity*: measures of neurological activity while participants processed information about a member of a majority or minority group; (b) *response time*: measures of stimulus response latencies, such as Correll's shooter task (Correll, Park, Judd, & Wittenbrink, 2002); (c) *microbehavior*: measures of nonverbal and subtle verbal behavior, such as displays of emotion and body posture during intergroup interactions and assessments of interaction quality based on reports of those interacting with the participant or coding of interactions by observers (this category encompasses behaviors Sue et al., 2007, characterized as "racial microaggressions"); (d) *interpersonal behavior*: measures of written or verbal behavior during an intergroup interaction or explicit expressions of preferences in an intergroup interaction, such as a choice in a Prisoner's Dilemma game or choice of a partner for a task; (e) *person perception*: explicit judgments about others, such as ratings of emotions displayed in the faces of minority or majority targets or ratings of academic ability; (f) *policy/political preferences*: expressions of preferences with respect to specific public policies that may affect the welfare of majority and minority groups (e.g., support for or opposition to affirmative action and deportation of illegal immigrants) and particular political candidates (e.g., votes for Obama or McCain in the 2008 presidential election).

These distinctions among criteria allow for tests of extant theory and also provide practical insights into the nature of IAT prediction. Many theorists have contended that implicit bias leads to discriminatory outcomes through its impact on microbehaviors that are expressed, for instance, during employment interviews and on quick, spontaneous reactions of the kind found in Correll's shooter task, and they contend that the effects of implicit bias are less likely to be found in the kind of deliberate choices involved in explicit personnel decisions (e.g., Chugh, 2004; Greenwald & Krieger, 2006; see Mitchell & Tetlock, 2006; Ziegert & Hanges, 2005). Furthermore, because the expression of political preferences can be easily justified on legitimate grounds that avoid attributions of prejudice (Sears & Henry, 2005), participants should be less motivated to control and conceal biased responding on the policy preference criteria, and we should thus find stronger correlations with implicit bias in this category (cf. Fazio, 1990; Olson & Fazio, 2009), and with explicit bias if it is measured in a way that reduces social desirability pressures on respondents (Sears, 2004b). Our criterion categories permit testing of these theoretical distinctions about the role of implicit and explicit bias for various kinds of prejudice and discrimination that have direct relevance for a broad range of theories.

Any attempt to reduce the diverse criteria found in the IAT studies to a single dimension of controllability would encounter the coding difficulties encountered by Greenwald, Poehlman, et al. (2009), while at the same time imposing arbitrary and potentially misleading distinctions. One crucial problem with such an approach is that post hoc judgments of the likely opportunity for psychological control available on a criterion task, even if those judgments are accurate, do not take into account the crucial additional factor of motivation to control prejudiced responses. Empirically supported theories of the relation of prejudicial attitudes to discrimination identify motivation as a key moderator variable in this relation (Dovidio et al., 2009; Olson & Fazio, 2009). Many of the IAT criterion studies did not include individual difference measures of motivation to control prejudice and neither manipulated nor measured felt motivation to avoid prejudicial responses. In short, we determined that coding criteria for opportunity and motivation to control responses could not be done in a reliable and meaningful way for the studies in our meta-analysis.

Nevertheless, the criterion categories we employ capture qualitative differences in participant behavior recorded by measures that may be a systematic source of variation in effects, and these qualitative differences can be leveraged to test competing theories of the nature of attitude–behavior relations. All of the criteria in the response time category involve tasks that permit little conscious control of behavior, and all of the criteria in the microbehavior category involve subtle aspects of behavior that were often measured unobtrusively. Both single-association models (which posit that implicit constructs bear the same relation to all forms of behavior) and double-dissociation models (which posit that implicit constructs have a greater influence on spontaneous behavior) predict that implicit bias should reliably predict behavior in these categories (see Perugini et al., 2010). Single-association models predict that implicit bias will also reliably predict more deliberative conduct. Thus, under the single-association view, implicit bias should also predict the explicit expressions of preferences, judgments, and choices found in the policy preferences, person perception, and interpersonal behavior criterion categories. Under double-dissociation models, explicit bias should be a stronger predictor of criteria found in the interpersonal behavior and person perception categories because they involve more deliberate action than criteria in the response time and microbehavior categories; race- or ethnicity-based distinctions will be harder to justify or deny in the tasks involved in those criterion categories compared to the policy preference category.[4] The research literature contains conflicting ev-

[4] We do not make a prediction for brain activity criteria because we do not consider neuroimages to be forms of behavior. We return to this point in the Discussion.

idence about the accuracy of single-association and double-dissociation models (Perugini et al., 2010). Our criterion-measure moderator analyses cannot precisely determine why some criteria are more or less subject to influence by implicit or explicit biases, but these analyses will provide important data for this ongoing debate.[5]

> *Nature of the IAT:* Are attitude and stereotype IATs equally predictive of discrimination?

We examined whether the nature of the IAT affected prediction, with effects coded as either based on an attitude IAT (which seeks to measure evaluative associations) or based on a stereotype IAT (which seeks to measure semantic associations). If attitude and stereotype IATs capture different types of associations that serve different appraisal and behavior-guiding functions (Greenwald & Banaji, 1995), then prediction for some criterion measures should be more sensitive to the semantic *content* of concept associations (as measured by stereotype IATs), and prediction within other criterion measures should be more sensitive to the *valence* of concept associations (as measured by attitude IATs; but see Talaska et al., 2008, who found that stereotypic beliefs were less predictive of discriminatory behavior than attitudes and emotional prejudice). We predicted that stereotype IATs would be more predictive than attitude IATs of judgments on person perception tasks, on the theory that semantic associations should be more correspondent to the attributional inferences that must be drawn in the appraisal processes associated with these tasks. We predicted that attitude IATs would be more predictive of policy preferences, on the theory that implicit prejudice toward minorities should be more correspondent with evaluations of specific candidates and policies that benefit or disadvantage minority groups (e.g., Greenwald, Smith, et al., 2009).[6] We cannot compare the predictive validity of stereotype and attitudes IATs for other categories of criterion measures, because so few studies used the stereotype IAT to predict other criteria.[7]

> *Relative versus absolute criterion scoring:* Does the relation of IAT scores and explicit measures to criteria vary as a function of the manner by which criteria are scored?

Criterion measures of discrimination are typically derived in one of three ways: (a) directly rating the majority and minority group targets relative to one another on the same measure (e.g., a relative rating of academic ability), (b) computing the difference score on separate ratings of majority and minority group targets on the same measure (e.g., seating distance from member of majority group minus seating distance from member of minority group), or (c) rating the majority or minority group targets separately on the same metric, with no comparison between targets (e.g., rating of majority or minority target's trustworthiness; often, only ratings for the minority target are collected and reported using this approach). We examine the impact of these three criterion scoring methods on predictive utility. This approach contrasts with Greenwald, Poehlman, et al. (2009), who aggregated effects across the scoring methods (and, in the race domain, they also excluded effects for criterion measures that had been scored only for majority targets).

Comparisons of predictive validity across absolute versus relative criterion coding has important applied implications: It speaks to the need to understand whether the IAT is equally predictive of favorable treatment of majority members versus unfavorable treatment of minority members but, most important, to show that the IAT is predictive of relative comparisons, which are needed to ensure that majority and minority members are being treated differently (i.e., some persons may treat all persons unfavorably, regardless of their race or ethnicity, which would not constitute discrimination). However, this analysis also has theoretical significance, as it bears on the construct validity of the IAT, which conceptualizes attitudes in relativistic terms (evaluations of Whites vs. Blacks or of insects vs. flowers). This measurement approach assumes that implicit attitudes can be validly measured through reactions to potentially opposing attitude objects (Schnabel, Asendorpf, & Greenwald, 2008). Arguing against this assumption, Pittinsky and colleagues found distinct behavioral effects for both positive and negative attitudes toward minorities (Pittinsky, 2010; Pittinsky, Rosenthal, & Montoya, 2011), and, similarly, we have found that IAT scores can differentially predict interactions with White versus Black persons (Blanton et al., 2009). These results illustrate the need to examine the utility of the IAT in the prediction of relative versus absolute treatment of majority and minority members, as a systematic review of the broader literature could provide a more definitive analysis of the viability of the IAT measurement strategy.

> *Nature of the explicit measure:* Are some explicit measures more predictive of discrimination than others?

The defining feature of the studies we examine is that each contains an IAT measure. Many studies also contained explicit attitude measures that pitted against IAT instruments for prediction purposes, and, as discussed above, Greenwald, Poehlman, et al. (2009) emphasized the performance of IATs relative to these

[5] Greenwald, Poehlman, et al. (2009) coded criterion measures for controllability and found no moderating effects on IAT prediction for this variable, which they took as evidence against double-dissociation theories of attitude–behavior relations. But, as noted above, their conclusion was based on a moderator analysis that collapsed across all nine criterion domains covered by their meta-analysis, and Greenwald, Poehlman, et al. acknowledged that their finding of a null result on controllability could be due to the high correlations of IAT scores with criteria rated high on controllability in the political and consumer preference domains. Social pressure is likely to be much greater with respect to the expression of ethnic and racial preferences than the expression of political and consumer preferences, and these domains therefore serve as better tests of double-dissociation models. Social norms against discrimination should motivate people to avoid expressions of bias in those behaviors when there is an opportunity to do so, such as on person perception tasks, but implicitly measured bias should still predict response times and microbehaviors in these domains (see Dovidio et al., 2009; Olson & Fazio, 2009).

[6] The partisan nature of the political topics makes this domain one in which researchers can measure evaluation of polarizing attitude objects (e.g., President Obama, Senator McCain), such that they have a reasonable expectation that favorable and unfavorable evaluations will be strongly and negatively correlated. Moreover, political choices used as criteria in these studies (e.g., voting for Obama or McCain) situate decisions around this attitude structure (see Greenwald, Smith, et al., 2009). These conditions are precisely the psychometric conditions identified by Blanton et al. (2006, 2007) as most conducive to increasing IAT prediction of criteria.

[7] One study meeting our inclusion criteria used a stereotype IAT to predict interpersonal behavior, one used a stereotype IAT to predict microbehavior, one to predict policy preferences, and one to predict response times.

explicit measures in the interracial and other intergroup relations domains. A closer examination of the explicit measures used in the IAT criterion studies is necessary to understand (a) whether the IAT performed better than all kinds of explicit measures across all forms of prejudice and discrimination and (b) why the validity levels reported by Greenwald, Poehlman, et al. for explicit measures were so much lower than the validity levels previously found for explicit measures of prejudice (Kraus, 1995; Talaska et al., 2008).

Ideally, we would have coded the explicit measures in IAT studies for the degree to which measurement was informed by lessons from modern attitude theory, and in particular whether researchers followed what Fishbein and Ajzen (2010) called the *principle of compatibility*: "an intention is compatible with a behavior if both are measured at the same level of generality or specificity—that is, if the measure of intention involves exactly the same action, target, context, and time elements as the measure of behavior" (Fishbein & Ajzen, 2010, p. 44). This issue is most relevant to studies that examine behavioral outcomes (i.e., microbehaviors, interpersonal behaviors), but studies focusing on other criterion domains could have been informed by this concern (see Jaccard & Blanton, 2006). Regardless of the criterion domain, researchers should strive to situate their explicit measures within what Cronbach and Meehl (1955) termed a nomological network, a causal framework that considers such issues as how compatibility might affect the degree of causal association that can be observed. For instance, in some of the person perception studies designed to predict judgments of guilt or innocence, researchers employed explicit measures that tapped general feelings toward members of different groups or distal beliefs related to the economic standing of their members (e.g., Florack, Scarabis, & Bless, 2001; Levinson et al., 2010). Had the principle of compatibility been given consideration, they might instead have pursued explicit measures designed to assess the perceived moral character or criminal nature of different group members. Generally, studies in the person perception domain could have taken advantage of the compatibility principle by tailoring their explicit measures to the goal of predicting employment, academic, or social outcomes.

We found little evidence that explicit measures were tailored to the criteria of interest and so were unable to code for compatibility (i.e., the measures would have been rated low in compatibility in almost all studies in our meta-analysis). It also would be ideal to code for the presence of efforts designed to minimize social desirability biases in explicit responses (see, e.g., Tourangeau & Yan, 2007), but there was insufficient reporting of steps taken to address this issue. We did, however, differentiate among the kinds of explicit measures used in the criterion studies to determine their levels of validity in predicting the various kinds of discrimination examined in these studies. In particular, we distinguished among (a) feeling thermometers, which assessed how warmly or coolly participants felt toward different groups, (b) established measures of bias that assess broad intergroup attitudes and stereotypes (e.g., the Modern Racism Scale, which was often used in the criterion studies synthesized here), and (c) ad hoc measures created for the individual study that typically involved one or a few questions aimed at gathering general attitudes toward or stereotypic beliefs about different groups. This differentiation did allow one prediction regarding compatibility: We predicted that established measures of bias would be more predictive of political preferences than of other criterion measures because of attitude–behavior compatibility at the level of public policy support (e.g., affirmative action) but not at the level of everyday interactions with particular individuals (e.g., making judgments about how happy, sad, or angry another person is). Overall, our inquiries into the validities across explicit measures are exploratory in nature and are aimed at confirming the low explicit-criterion correlations reported by Greenwald, Poehlman, et al. (2009).

Methodological Refinements

Meta-analytic approach. This study demonstrates how to deal meta-analytically with multiple-effect sizes when they represent different operationalizations of the same general construct, such as acts of discrimination, and are derived from a common sample. (For discussions of the issues presented by such studies and different methodological options, see Cheung & Chan, 2004, 2008; Gleser & Olkin, 2007; Hedges, Tipton, & Johnson, 2010; Kim & Becker, 2010).[8] In many meta-analyses in the social sciences, relatively few studies involve multiple criteria or multiple behavioral outcomes from the same sample, making the treatment of dependent effects relatively unimportant in statistical estimation. This issue is pivotal, however, in any empirical examination of the predictive utility of a psychological inventory that is designed for use in studies that predict a broad array of psychological criteria, across a wide range of social contexts—a description that applies to a large number of the IAT criterion studies. We thus employed the random-effects model of meta-analysis proposed by Hedges et al. (2010), which incorporates the dependence between multiple correlations drawn from the same sample. Their method deals parsimoniously with the heterogeneity of effects within studies containing multiple effects and does not require knowledge of the sampling distributions of the effects. Instead, a single parameter estimate for the correlation between all dependent correlations is used. We assumed this value to be $r = .50$, as recommended by the developers of this model, but as the developers also recommend, we conducted follow-up analyses that varied this value. Only trivial changes in results and interpretation were found.

Our meta-analytic approach contrasts with that of Greenwald, Poehlman, et al. (2009), who, as noted, averaged across multiple criteria to produce a single effect size estimate for each sample. Across all 184 IAT-criterion effects used in Greenwald, Poehlman, et al., 44 were based on averaging 3 or more criteria within a sample, and 6 were based on averaging 10 or more criteria (with three of these six coming from the race domain). That approach necessarily reduces effect-size variability, leading to downwardly biased estimates, and, conceptually, that approach obscures substantive differences among criteria that the researchers in each study had originally set out to distinguish and investigate. Effect-size variability within samples (across dependent effects) is every bit as theoretically important to quantify and account for as variability between samples. The approach we pursue takes this variability into account while at the same time allowing flexibility in

[8] We appreciate the helpful input and meta-analysis expertise of Dan Beal, Mike Brannick, Mike Cheung, Ron Landis, and Scott Morris and the sharing of meta-analysis computer code and related support by Elizabeth Tipton.

categorizing effects within the same sample across different categories of moderators (i.e., this approach allows us to assign criteria to different moderator categories and still model effect-size dependencies, whereas Greenwald, Poehlman, et al. assigned all criteria within a sample a single value on a moderator variable and ignored dependencies).

Correcting data errors and omissions. Finally, the present meta-analysis corrects a number of errors found in the data set used in Greenwald, Poehlman, et al. (2009).[9] We report these corrections in detail in online supplemental materials. These errors influence the meta-analytic effect sizes and moderator analyses, and correction of these errors thus constitutes an important contribution in its own right, given the importance of the Greenwald, Poehlman, et al. meta-analysis within implicit social cognition research and applications of this research to the law and public policy.

Method

Literature Search and Inclusion Criteria

We supplemented the effects located by Greenwald, Poehlman, et al. (2009) for the racial and ethnic domains by (a) searching PsycINFO for post-2006 studies using the same search terms used by Greenwald, Poehlman, et al., (b) searching the Social Science Research Network (SSRN) for articles with the word *implicit* in the title or abstract, (c) requesting any in-press and unpublished studies examining the predictive validity of the IAT on the Society of Personality and Social Psychology(SPSP), JDM-Society, Social Psychology Network, CogSci, Neuro-psych, and Sociology & Psychology listservs (we fashioned our postings after the e-mail request made by Greenwald, Poehlman, et al. to the SPSP mailing list), and (d) examining the post-February 2007 studies collected by Greenwald, Poehlman, et al. that were made available in their online archive supporting their published article. We included any study for which an IAT-criterion correlation (an "ICC" to use Greenwald, Poehlman, et al.'s nomenclature) could be computed where the criterion arguably measured some form of discrimination. A number of studies computed correlations between IAT scores and responses to surveys that measured general attitudes or views about socioeconomic conditions; we excluded these studies on grounds that these correlations constituted implicit–explicit correlations (an "IEC" to use Greenwald, Poehlman, et al.'s nomenclature) rather than ICCs. We included studies in which IAT scores were correlated with responses to specific policy proposals, such as affirmative action and immigration policy. Our searches and inclusion criteria resulted in 46 published and unpublished reports of 308 ICCs and 275 explicit-criterion correlations ("ECCs" to use Greenwald, Poehlman, et al.'s nomenclature) based on 86 different samples.[10]

Moderator Variables

The second and third authors coded all studies for five moderator variables: (a) criterion domain (race or ethnicity/national origin), (b) IAT version (attitude/evaluative associations, stereotype/semantic associations, or other), (c) explicit measure utilized (preexisting measure of bias other than feeling thermometer, feeling thermometer, or measure created for the study), (d) criterion operationalization (interpersonal behavior, person perception, policy preference, microbehavior, response time, brain activity), and (e) criterion measure scoring method (relative rating of majority and minority group members on the same measure, difference score computed from ratings of majority and minority group members on common measure, or rating of a single group on the measure). This coding approach called for little subjective judgment, and there was very high agreement between coders, with the few initial disparities in coding reflecting simple mistakes that were easily resolved.

Calculation of Effect Sizes and Statistical Approach

To provide a descriptive sense of the typical effect size and variability observed within each meta-analysis conducted, the rightmost columns of our tables show unweighted means and standard deviations of the effect sizes for each meta-analysis.[11] We then report meta-analytic means and standard deviations (the latter being tau, an estimate of random effects) that are based on effect-size weights that are estimated iteratively and take into account the fact that some effects have more stable estimates than others by virtue of larger sample sizes. Dependencies between correlations from the same sample are also taken into account during this procedure, regardless of whether those correlations fall within the same category of moderator variable. Data for all meta-analytic results are available in the online supplemental materials. The R code program we applied is provided in Hedges et al. (2010).

Meta-analyses were conducted across levels of the moderator variables to obtain meta-analytic correlations and confidence intervals for each level, as well as the estimate of random-effects variation (tau) across levels after accounting for and reporting random-effects variation within each level. The corresponding unweighted mean and standard deviation were computed with a simple spreadsheet program. Tabled results provide the number of effects, number of independent samples, and the cumulative sample size (k, s, and N_{tot}, respectively) for each separate meta-analysis. Meta-analyses with small numbers of studies and effects are provided in the tables to provide more comprehensive description of the extant IAT literature, but these results should be interpreted cautiously for both statistical

[9] Our corrections go beyond incorporating a few studies inadvertently omitted from Greenwald, Poehlman, et al. (2009), though we do correct such omissions. In particular, we document (a) inconsistent requests by Greenwald, Poehlman, et al. regarding unpublished data, (b) inconsistent treatment of effects from studies with identical or nearly identical designs, (c) use of unclear or ad hoc criteria to select among available effects for inclusion in the data, (d) omission or exclusion of effects and studies that met inclusion criteria, (e) inconsistent coding of moderator variables, (f) inclusion of an erroneously reported effect, and (g) inclusion of an effect based on fabricated data. We describe these issues in detail in the online supplemental materials.

[10] Greenwald, Poehlman, et al. (2009) reported only 47 effects for the race and other intergroup criterion domains. The large difference in our number of effects is due to Greenwald, Poehlman, et al. averaging effects across conditions and using only a single mean effect per study. In addition, recall that we focused on racial and ethnic/national-origins biases and excluded studies of bias against religious groups, obese persons, and older persons, which Greenwald, Poehlman, et al. included in their "other intergroup relations" domain. Also, we included several articles on racial and ethnic bias that were published after Greenwald, Poehlman, et al.'s cutoff date.

[11] All effects were coded such that positive signs reflect promajority group or antiminority group bias scores on the IAT or explicit measures and higher discrimination scores on the criterion measures.

reasons (e.g., less precision for the estimates) and substantive reasons (e.g., concerns about the limited representativeness of the collection of studies). We report the standard deviations of the random effects (taus), but it is important to keep in mind that the standard error of tau or any estimate of heterogeneity can be large (Biggerstaff & Tweedie, 1997), especially when the number of studies is small (Borenstein, Hedges, Higgins, & Rothstein, 2009, p. 364; Hartung & Knapp, 2003; Oswald & Johnson, 1998). Reporting both the weighted mean and standard deviation from the meta-analysis for each set of effects, along with the unweighted mean and standard deviation, provides a comprehensive picture of the available empirical evidence.

Results

IAT Criterion-Related Correlations (ICCs)

Table 1 shows the meta-analytic estimates of criterion-related validities within and across the race and ethnicity domains for IATs. The average ICCs for each domain and for the combined domains are small ($\hat{\rho}$ = .14 overall, .15 for the race domain, and .12 for the ethnicity domain) and are empirically heterogeneous (e.g., both taus and the unweighted standard deviations are in fact equal to or larger than the corresponding estimated mean effects). Some of this heterogeneity is explained by differences in criterion measures: A much higher average correlation was found in the brain activity subdomain; the IAT did not predict microbehaviors well and no better than explicit measures for interpersonal behaviors, person perceptions, policy preferences, and response times.

Contrary to our prediction, predictive validities were particularly low for stereotype IATs on person perception tasks; however, stereotype IATs were used much less often than attitude IATs, and neither version of the IAT was a good predictor of discriminatory behavior, judgments, or decisions (see Table 2). Thus, the large amounts of heterogeneity reported in Table 1 are not explained by differences in the predictive validities of the attitude and stereotype IATs.

The predictive validities of IATs across criterion scoring methods are presented in Tables 3 and 4. Because these analyses subdivide ICC validities further—by criterion domain and then by criterion-scoring method—some of these findings are based on small numbers of effects and should be viewed as exploratory in nature. In the race domain, there is no clear pattern across criterion categories, but the data suggest that the IAT is generally a better predictor of behavior toward Black targets than White targets, because the ICCs tend to be larger for scoring methods that incorporate behavior directed at Black targets (i.e., difference scores and ratings of the Black target alone). Given that most of the samples were predominantly White, this finding could indicate that interracial attitudes were not activated to the same extent in same-race interactions. Further studies should explore the reliability and source of this

Table 1
Meta-Analysis of Implicit-Criterion Correlations (ICCs): Overall and by Subgroups

Criterion	k (s; N_{total})	$\hat{\rho}$ [95% CI]	$\hat{\tau}$	M	SD
All effects: Overall	298 (86; 17,470)	.14 [.10, .19]	.17	.12	.24
Interpersonal behavior	11 (6; 796)	.14 [.03, .26]	.12	.21	.15
Person perception	138 (46; 7,371)	.13 [.07, .18]	.13	.10	.21
Policy preference	21 (9; 4,677)	.13 [.07, .19]	.03	.14	.09
Microbehavior$_a$	96 (21; 3,879)	.07 [–.03, .18]	.19	.10	.24
Response time	6 (5; 300)	.19 [.02, .36]	.27	.31	.28
Brain activity$_a$	26 (8; 447)	.42 [.11, .73]	.68[b]	.26	.40
Black vs. White groups: Overall	206 (63; 9,899)	.15 [.09, .21]	.19	.13	.26
Interpersonal behavior	10 (5; 691)	.14 [.01, .28]	.14	.22	.16
Person perception	75 (30; 3,564)	.13 [.08, .19]	.12	.09	.22
Policy preference$_a$	8 (5; 1,855)	.10 [.02, .19]	.05	.09	.10
Microbehavior$_b$	87 (18; 3,162)	.07 [–.06, .19]	.22	.10	.25
Response time[a]	6 (5; 300)	.19 [.02, .37]	.27	.31	.28
Brain activity$_{a,b}$	20 (8; 327)	.43 [.12, .73]	.67[b]	.30	.42
Ethnic minority vs. majority groups: Overall	92 (24; 7,571)	.12 [.06, .19]	.12	.12	.18
Interpersonal behavior$_a$	1 (1; 105)	.19 [[c]]	—	.19	[c]
Person perception	63 (16; 3,807)	.11 [–.01, .23]	.15	.11	.19
Policy preference	13 (4; 2,822)	.16 [.08, .25]	.00	.17	.07
Microbehavior	9 (3; 717)	.11 [–.09, .31]	.14	.11	.19
Response time	—	—	—	—	—
Brain activity$_a$	6 (1; 120)	.11 [[c]]	—	.11	.27

Note. All effects were coded such that positive correlations are in the direction of promajority group or antiminority group responses or behaviors. With regard to Heider and Skowronski (2007) and Stanley et al. (2011), these analyses incorporate the difference score ICCs, not the Black-only and White-only ICCs. The correlation between dependent effects is assumed to be .50. The $\hat{\rho}$ for each category is based on a moderated meta-analysis across categories, where dependent effect sizes (both within and across categories) are accounted for (Hedges, Tipton, & Johnson, 2010), and the overall random-effects variance (tau-squared) weight is applied. $\hat{\tau}$ is also independently estimated within each category in separate analyses. Dashes indicate insufficient number of effects for computation purposes. Effects sharing subscripts within a category set are statistically significantly different from one another ($p < .05$). k = number of effects; s = number of independent samples within each category (this does not add up to the overall s because of sample overlap across categories); $\hat{\rho}$ = meta-analytically estimated population correlation; CI = confidence interval; $\hat{\tau}$ = random-effects standard deviation estimate; M = unweighted mean; SD = unweighted standard deviation.
[a] Even though this category in the overall analysis and in the Black-only analysis contains the same effects, results differ because estimates within categories are influenced by the effects, dependencies, and weighting across categories. [b] This extremely large value is in fact the estimated value. [c] An appropriate estimate cannot be computed due to the integrated analysis with limited effects in this category.

Table 2
Meta-Analysis of Attitude and Stereotype ICCs: Person Perception Criterion

IAT type	k (s; N_{total})	$\hat{\rho}$ [95% CI]	$\hat{\tau}$	M	SD
All effects					
Attitude IAT_{a}	97 (40; 5,096)	.16 [.11, .21]	.10	.09	.19
Stereotype IAT_{a}	41 (14; 2,275)	.03 [–.08, .14]	.16	.12	.23
Black vs. White					
Attitude IAT_{a}	51 (26; 2,627)	.17 [.11, .23]	.11	.10	.20
Stereotype IAT_{a}	24 (10; 937)	.06 [–.02, .14]	.15	.06	.25
Asian vs. non-Asian					
Attitude IAT	15 (2; 1,243)	.04 [.03, .04]	.00	.04	.12
Stereotype IAT	17 (4; 1,338)	–.03 [–.82, .76]	.21	.21	.18

Note. The comparisons reported are limited to those with larger numbers of effects. All effects were coded such that positive correlations are in the direction of promajority group or antiminority group responses or behaviors. The correlation between dependent effects is assumed to be .50. The $\hat{\rho}$ for each category is based on a moderated meta-analysis across categories, where dependent effect sizes (both within and across categories) are accounted for (Hedges et al., 2010) and Stanley et al. (2011), and the overall random-effects variance (tau-squared) weight is applied. $\hat{\tau}$ is also independently estimated within each category in separate analyses. With regard to Heider and Skowronski (2007), these analyses incorporate the difference score ICC, not the Black-only and White-only ICCs. Effects sharing subscripts within a category set are significantly different from one another ($p < .05$). ICCs = implicit-criterion correlations; k = number of effects; s = number of independent samples within each category (this does not add up to the overall s because of sample overlap across categories); $\hat{\rho}$ = meta-analytically estimated population correlation; CI = confidence interval; $\hat{\tau}$ = random-effects standard deviation estimate; M = unweighted mean; SD = unweighted standard deviation; IAT = Implicit Association Test.

difference in predictive validity for same-race and different-race interactions. Response-time difference scores and response times to Black targets correlated moderately to highly with IAT scores, suggesting either shared method variance or a similarity in the relative mental comparisons required by the IAT and criterion tasks; however, these correlations are based on a small number of effects.

In the ethnicity domain, only person perception tasks contained sufficient variation in scoring methods to allow comparisons. The IAT was equally predictive of ratings of minority and majority targets on person perception tasks but, interestingly, was a very weak predictor of difference scores on person perception tasks (and note that the estimate for the difference-scored person perception criterion is based on a large number of effects relative to many of the other estimates in this domain). These scoring method analyses demonstrate not only tremendous heterogeneity in ICCs across studies but also considerable diversity in the research designs used in the IAT criterion studies and the lack of a common measurement procedure across studies (see De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009, on the need for greater standardization).

Explicit Measure Criterion-Related Correlations (ECCs)

Table 5 reports the meta-analytic estimates of criterion-related validities within and across the race and ethnicity domains for explicit measures of bias. ECCs overall and within the race and ethnicity domains were small and similar in magnitude to the ICCs ($\hat{\rho}$ = .12 overall, .10 for race, and .15 for ethnicity). However, the explicit measures showed greater variation in meta-analytic validity than the IATs across criterion operationalizations. Explicit measures were poor predictors of microbehavior in both the race and ethnicity domains ($\hat{\rho}$ = .02 and .11, respectively); they were somewhat better predictors of interpersonal behavior, policy preferences, and response times. Explicit measures were most predictive of brain activity, though only two effects from the race domain were available ($\hat{\rho}$ = .33) on the brain-activity criterion. Notably, the validities for interpersonal behavior and person perception—criterion categories involving more controlled behavior and having the most direct connection to discriminatory behavior—were generally as low as those for the IAT, the highest being $\hat{\rho}$ = .19 for interpersonal behavior in the race domain based on eight effect sizes (the person perception effects, which were based on more effects, were both lower than this value, $\hat{\rho}$s = .11).

The method by which explicit constructs were assessed made little difference (see Table 6, bottom section), as the low predictive validities held in both race and ethnicity domains for feeling thermometers, preexisting bias scales, and ad hoc measures (although in the ethnicity domain, ad hoc measures showed higher validity). In general, the explicit measures performed below the level one would expect for simple, general attitude measures used to predict specific behaviors (see Wicker, 1969) and below the levels found previously for explicit prejudice–behavior relations (Kraus, 1995; Talaska et al., 2008). This suggests that greater attention to strategies for improving explicit measurement might improve their performance relative to implicit measures (see, e.g., Ditonto, Lau, & Sears, in press).

Tables 7 and 8 provide ECCs across the criterion-scoring methods and criterion measure categories. In the race domain, explicit measures were better predictors of interpersonal behavior and person perception when they consisted of ratings of Black targets or used relative ratings. In the ethnicity domain, only three studies report effects for criterion measures scored for majority targets only. Several studies in the ethnicity domain report effects for ratings of minority targets only, without corresponding ratings of majority targets, making it difficult to assess whether any disparate treatment occurred within these studies (see Blanton & Mitchell, 2011).

Table 3
Meta-Analysis of ICCs by Criterion Scoring Method: Black Versus White Groups

Criterion scoring method	k (s; N_{total})	$\hat{\rho}$ [95% CI]	$\hat{\tau}$	M	SD
Overall					
Absolute—Black target$_a$	65 (27; 3,601)	.15 [.06, .23]	.17	.15	.25
Absolute—White target$_{a,b}$	33 (19; 1,344)	–.01 [–.07, .05]	.07	.01	.17
Relative rating	14 (10; 1,496)	.13 [–.01, .26]	.12	.09	.24
Difference score$_b$	104 (26; 4,144)	.22 [.10, .34]	.33[a]	.15	.28
Interpersonal behavior					
Absolute—Black target$_a$	9 (4; 628)	.23 [.19, .27]	.00	.25	.09
Absolute—White target$_a$	3 (3; 246)	–.13 [−.24, −.03]	.00	–.11	.07
Relative rating	—	—	—	—	—
Difference score	2 (2; 183)	.14 [−.19, .48]	.24	.22	.27
Person perception					
Absolute—Black target	35 (11; 1,816)	.19 [.07, .31]	.19	.12	.24
Absolute—White target	18 (9; 802)	.03 [–.08, .14]	.03	–.02	.15
Relative rating	9 (7; 359)	.15 [.08, .22]	.00	.14	.15
Difference score	15 (8; 687)	.19 [.07, .31]	.14	.12	.24
Policy preference					
Absolute—Black target	7 (4; 798)	.08 [–.10, .26]	.03	.08	.10
Absolute—White target	—	—	—	—	—
Relative rating	1 (1; 1,057)	.17 [[b]]	—	.17	—
Difference score	—	—	—	—	—
Microbehavior					
Absolute—Black target	9 (6; 278)	.00 [–.33, .33]	.27	.00	.33
Absolute—White target	7 (5; 215)	–.02 [–.20, .16]	.14	.04	.25
Relative rating	4 (3; 80)	.04 [–.54, .62]	.48[a]	−.02	.42
Difference score	71 (8; 2,809)	.14 [.04, .25]	.12	.12	.22
Response time					
Absolute—Black target$_a$	1 (1; 21)	.52 [[b]]	—	.52	—
Absolute—White target$_a$	1 (1; 21)	.06 [[b]]	—	.06	—
Relative rating	—	—	—	—	—
Difference score	4 (3; 258)	.32 [–.70, 1.00]	.31[a]	.32	.30
Brain activity					
Absolute—Black target$_a$	4 (2; 60)	.54 [.41, .68]	.00	.54	.11
Absolute—White target$_a$	4 (2; 60)	.13 [–.04, .29]	.00	.12	.17
Relative rating	—	—	—	—	—
Difference score	12 (6; 207)	.36 [–.35, 1.00]	.73[a]	.28	.51

Note. All effects were coded such that positive correlations are in the direction of promajority group or antiminority group responses or behaviors. The correlation between dependent effects is assumed to be .50. The $\hat{\rho}$ for each category is based on a moderated meta-analysis across categories, where dependent effect sizes (both within and across categories) are accounted for (Hedges et al., 2010), and the overall random-effects variance (tau-squared) weight is applied. $\hat{\tau}$ is also independently estimated within each category in separate analyses. Estimated confidence interval bounds with magnitudes exceeding 1.00 were truncated at 1.00. Dashes indicate insufficient number of effects for computation purposes. Effects sharing subscripts within a category set are statistically significantly different from one another ($p < .05$). ICCs = implicit-criterion correlations; k = number of effects; s = number of independent samples within each category (this does not add up to the overall s because of sample overlap across categories); $\hat{\rho}$ = meta-analytically estimated population correlation; CI = confidence interval; $\hat{\tau}$ = random-effects standard deviation estimate; M = unweighted mean; SD = unweighted standard deviation.
[a] This extremely large value is in fact the estimated value. [b] An appropriate estimate cannot be computed due to the integrated analysis with limited effects in this category.

Correlations Between IATs and Explicit Measures (IECs)

In the race domain, meta-analytically averaged correlations between implicit and explicit measures were low overall ($\hat{\rho}$ = .14), with IECs for preexisting measures and ad hoc measures higher than for feeling thermometers (see Table 6). A higher IEC was found in the ethnicity domain for feeling thermometers, but IECs were somewhat lower for ad hoc measures. The low IECs found in the race domain are comparable to those found by Greenwald, Poehlman, et al. (2009) but lower than that found by Nosek et al. (2007; r = .27) based on an extremely large web-based sample (N = 732,881). These findings collectively indicate, at least for the race domain, either that implicit and explicit measures tap into different psychological constructs—none of which may have much influence on behavior, given the low ICCs and ECCs observed—or that social or methodological factors adversely affect the validity of responses to explicit measures of racial bias and possibly the race IAT as well (e.g., Frantz, Cuddy, Burnett, Ray, & Hart, 2004).[12] Greenwald, Poehlman, et al. (2009) theorized that reactive measurement effects and the limits of introspection adversely

[12] The fact that IATs and explicit measures converged in their predictions of brain activity might be seen as counterevidence in favor of the view that both types of measures tap into constructs that are associated with the same brain processes activated in interracial and interethnic interactions. But this evidence of convergence should be viewed as tentative, given the small number of studies and sample sizes on which the effects are based and given problems with the reporting of results from these studies (see online supplemental materials; see also Vul, Harris, Winkielman, & Pashler, 2009).

Table 4
Meta-Analysis of ICCs by Criterion Scoring Method: Ethnic Minority Versus Majority Groups

Criterion scoring method	k (s; N_{total})	$\hat{\rho}$ [95% CI]	$\hat{\tau}$	*M*	*SD*
Overall					
Absolute—Minority target	29 (12; 3,614)	.16 [.09, .23]	.08	.11	.17
Absolute—Majority target	11 (4; 510)	.18 [.05, .31]	.05	.11	.15
Relative rating	—	—	—	—	—
Difference score	52 (10; 3,447)	.07 [–.07, .20]	.17	.13	.20
Interpersonal behavior					
Absolute—Minority target	1 (1; 105)	.19 [a]	—	.19	—
Absolute—Majority target	—	—	—	—	—
Relative rating	—	—	—	—	—
Difference score	—	—	—	—	—
Person perception					
Absolute—Minority target	14 (7; 582)	.16 [.00, .33]	.18	.05	.23
Absolute—Majority target	11 (4; 510)	.18 [.04, .32]	.05	.11	.15
Relative rating	—	—	—	—	—
Difference score	38 (7; 2,715)	.05 [–.14, .23]	.18	.14	.19
Policy preference					
Absolute—Minority target	13 (4; 2,822)	.17 [.08, .25]	.00	.17	.07
Absolute—Majority target	—	—	—	—	—
Relative rating	—	—	—	—	—
Difference score	—	—	—	—	—
Microbehavior					
Absolute—Minority target	1 (1; 105)	.08 [a]	—	.08	—
Absolute—Majority target	—	—	—	—	—
Relative rating	—	—	—	—	—
Difference score	8 (2; 612)	.12 [a]	.23	.11	.20
Response time					
Absolute—Minority target	—	—	—	—	—
Absolute—Majority target	—	—	—	—	—
Relative rating	—	—	—	—	—
Difference score	—	—	—	—	—
Brain activity					
Absolute—Minority target	—	—	—	—	—
Absolute—Majority target	—	—	—	—	—
Relative rating	—	—	—	—	—
Difference score	6 (1; 120)	.11 [a]	—	.11	.27

Note. All effects were coded such that positive correlations are in the direction of promajority group or antiminority group responses or behaviors. The correlation between dependent effects is assumed to be .50. The $\hat{\rho}$ for each category is based on a moderated meta-analysis across categories, where dependent effect sizes (both within and across categories) are accounted for (Hedges et al., 2010), and the overall random-effects variance (tau-squared) weight is applied. $\hat{\tau}$ is also independently estimated within each category in separate analyses. Dashes indicate insufficient number of effects for computation purposes. Estimated confidence interval bounds with magnitudes exceeding 1.00 were truncated at 1.00. No effects within a category set are statistically significantly different from one another ($p < .05$). ICCs = implicit-criterion correlations; k = number of effects; s = number of independent samples within each category (this does not add up to the overall s because of sample overlap across categories); $\hat{\rho}$ = meta-analytically estimated population correlation; CI = confidence interval; $\hat{\tau}$ = random-effects standard deviation estimate; M = unweighted mean; SD = unweighted standard deviation.
[a] An appropriate estimate cannot be computed due to the integrated analysis with limited effects in this category.

affect the utility of explicit measures in socially sensitive domains. Our results suggest that a finer grained approach should be taken, one that examines the sensitive nature of the topics, the particular measures used, and the efforts made to reduce social desirability pressures (e.g., some people may be much more comfortable expressing negative attitudes toward illegal immigrants than against African Americans, particularly if doing so in a setting that ensures anonymity or frames the topic as a matter of legitimate political debate; cf. Hofmann et al., 2005).

Incremental Validity Analysis

Greenwald, Poehlman, et al. (2009) used the meta-analytic correlations for ICCs, ECCs, and IECs to calculate rough estimates of incremental gain; namely, how much variance the IAT measure predicts over and above explicit measures and vice versa. We conducted a similar analysis (see Table 9). In light of the low magnitudes of the ICCs and ECCs, it is not surprising that the percentage of criterion variance they account for jointly is small (endpoints ranging from 2.4 to 3.2% for race and 1.6 to 6.8% for ethnicity) and that the amounts of incremental variance of ICCs over ECCs, and vice versa, were small (endpoints ranging from 0.1 to 2.0% for race and 0.2 to 5.4% for ethnicity).

Outlier Analysis

This meta-analysis included one large-N study that could dominate weighted estimates of average correlations; namely, the study

Table 5
Meta-Analysis of Explicit-Criterion Correlations (ECCs): Overall and by Subgroups

Criterion	k (s; N_{total})	$\hat{\rho}$ [95% CI]	$\hat{\tau}$	M	SD
All effects: Overall	263 (64; 18,223)	.12 [.07, .16]	.15	.08	.19
Interpersonal behavior	9 (3; 769)	.19 [–.03, .41]	.18	.28	.20
Person perception	124 (34; 5,797)	.11 [.03, .19]	.16	.09	.19
Policy preference$_a$	31 (8; 7,480)	.16 [.07, .25]	.16	.14	.17
Microbehavior$_{a,b}$	92 (18; 3,868)	.04 [–.04, .11]	.18	.02	.17
Response time$_b$	5 (4; 284)	.22 [.14, .31]	.02	.26	.14
Brain activity	2 (2; 25)	.34 [.02, .66]	.20	.28	.33
Black vs. White groups: Overall	198 (47; 12,706)	.10 [.05, .16]	.16	.07	.19
Interpersonal behavior	8 (2; 664)	.19 [–.09, .48]	.27	.29	.21
Person perception	79 (23; 3,445)	.11 [.01, .20]	.16	.07	.18
Policy preference	21 (5; 5,137)	.11 [–.02, .25]	.22	.10	.19
Microbehavior$_a$	83 (15; 3,151)	.02 [–.06, .09]	.04	.02	.17
Response time$_a$[a]	5 (4; 284)	.23 [.13, .32]	.02	.26	.14
Brain activity[a]	2 (2; 25)	.33 [–.00, .66]	.20	.28	.33
Ethnic minority vs. majority groups: Overall	65 (17; 5,517)	.15 [.05, .24]	.15	.13	.19
Interpersonal behavior	1 (1; 105)	.18 [[b]]	—	.18	—
Person perception	45 (11; 2,352)	.11 [–.06, .29]	.20	.12	.21
Policy preference	10 (3; 2,343)	.23 [.17, .29]	.00	.22	.07
Microbehavior	9 (3; 717)	.11 [–.14, .36]	.28	.06	.17
Response time	—	—	—	—	—
Brain activity	—	—	—	—	—

Note. All effects were coded such that positive correlations are in the direction of promajority group or antiminority group responses or behaviors. The correlation between dependent effects is assumed to be .50. The $\hat{\rho}$ for each category is based on a moderated meta-analysis across categories, where dependent effect sizes (both within and across categories) are accounted for (Hedges et al., 2010), and the overall random-effects variance (tau-squared) weight is applied. $\hat{\tau}$ is also independently estimated within each category in separate analyses. Dashes indicate insufficient number of effects for computation purposes. Effects sharing subscripts within a category set are statistically significantly different from one another ($p < .05$). k = number of effects; s = number of independent samples within each category (this does not add up to the overall s because of sample overlap across categories); $\hat{\rho}$ = meta-analytically estimated population correlation; CI = confidence interval; $\hat{\tau}$ = random-effects standard deviation estimate; M = unweighted mean; SD = unweighted standard deviation.
[a] Although this category contains the same data as the overall analysis, results can differ because estimates within categories are influenced by the effects, dependencies, and weighting across categories. With regard to Heider and Skowronski (2007) and Stanley et al. (2011), these analyses incorporate only the difference score ECCs. [b] An appropriate estimate cannot be computed due to the integrated analysis with limited effects in this category.

by Greenwald, Smith, et al. (2009). This study contributed seven effects: one ICC effect, three ECC effects, and three IEC effects, each with $N = 1{,}057$. To put this sample size in the context of the entire body of IAT research that we meta-analyzed, the next highest sample size was $N = 333$, and the median sample sizes for the overall ICC, ECC, and IEC were much smaller: 41, 41, and 77, respectively. We investigated whether this large-sample study altered in a nontrivial way any of the key parameters we report, and we generally did not find this to be the case. At the overall level of analysis, all meta-analytic correlations remained within .02 correlation units of the original estimate when Greenwald, Smith, et al. (2009) effects were excluded except that (a) the ICC for policy preferences within the race domain changed from .10 to .07 (see Table 1) and (b) the ECC for policy preference within the race domain changed from .11 to .06 (see Table 5). Note that we also provide unweighted means and standard deviations, which do not favor large-sample effects and which provide a similar pattern of effects as the sampling-error and random-effects variance-weighted counterparts from meta-analysis.

Discussion

Our meta-analytic estimates of the mean correlations between IAT scores and criterion measures of racial and ethnic discrimination are smaller than analogous correlations reported by Greenwald, Poehlman, et al. (2009): overall correlations of .15 and .12 for racial and interethnic behavior compared to correlations of .24 and .20 for racial and other intergroup behavior reported by Greenwald and colleagues. We arrived at different estimates for two reasons. First, Greenwald, Poehlman, et al. averaged multiple effects that were dependent on the same sample. Although this is not an uncommon practice, it suppresses substantive variance in the service of meeting the assumption of independent effects in a typical random-effects meta-analysis model. Second, we included a number of effects that were not available to Greenwald, Poehlman, et al., and we included effects erroneously omitted or erroneously coded in their earlier meta-analysis (see the online supplemental materials for details). Many of these additions involved weaker correlations between IAT scores and criterion measures.

The focused analysis of IAT–criterion correlations by the nature of the criterion measure also revealed that the validity estimates provided by Greenwald, Poehlman, et al. (2009) for the interracial and other intergroup relations domains appear to have been biased upward by effects from neuroimaging studies. IAT scores correlated strongly with measures of brain activity but relatively weakly with all other criterion measures in the race domain and weakly with all criterion measures in the ethnicity domain. IATs, whether they were designed to tap into implicit prejudice or implicit stereotypes, were typically poor predictors of the types of behavior, judgments, or decisions that have been studied as instances of

Table 6
Meta-Analysis of Implicit–Explicit Correlations (IECs) and Explicit-Criterion Correlations (ECCs) by Explicit Measure

Explicit measure	k (s; N_{total})	$\hat{\rho}$ [95% CI]	$\hat{\tau}$	M	SD
		IEC			
Black vs. White groups: Overall	105 (39; 10,739)	.14 [.09, .19]	.13	.13	.15
Thermometer	24 (15; 2,534)	.09 [–.05, .24]	.17	.14	.19
Other existing measure	39 (26; 3,491)	.15 [.09, .21]	.08	.16	.14
Created measure	42 (10; 4,714)	.14 [.05, .24]	.15	.10	.14
Ethnic minority vs. majority groups: Overall	19 (12; 1,339)	.16 [.09, .23]	.00	.13	.14
Thermometer_a	3 (2; 511)	.23 [.19, .27]	.08	.31	.11
Other existing measure	6 (5; 402)	.13 [.02, .24]	.00	.13	.09
Created measure_a	10 (7; 426)	.07 [–.03, .17]	.00	.07	.12
		ECC			
Black vs. White groups					
Thermometer	29 (18; 2,249)	.11 [–.05, .27]	.16	.07	.24
Other existing measure	112 (31; 6,008)	.11 [.05, .18]	.20	.06	.20
Created measure	57 (14; 4,449)	.06 [.00, .13]	.06	.07	.15
Ethnic minority vs. majority groups					
Thermometer	10 (5; 2,187)	.06 [–.16, .28]	.16	.14	.18
Other existing measure	14 (6; 917)	.09 [–.03, .22]	.04	.05	.10
Created measure	41 (8; 2,413)	.24 [.09, .40]	.18	.15	.21

Note. All effects were coded such that positive correlations are in the direction of promajority group or antiminority group responses or behaviors. The correlation between dependent effects is assumed to be .50. The $\hat{\rho}$ for each category is based on a moderated meta-analysis across categories, where dependent effect sizes (both within and across categories) are accounted for (Hedges et al., 2010), and the overall random-effects variance (tau-squared) weight is applied. $\hat{\tau}$ is also independently estimated within each category in separate analyses. With regard to Heider and Skowronski (2007), these analyses incorporate only the difference score ECCs. Effects sharing subscripts within a category set are statistically significantly different from one another ($p <$ .05). k = number of effects; s = number of independent samples within each category (may not add up to the overall s because of sample overlap across categories); $\hat{\rho}$ = meta-analytically estimated population correlation; CI = confidence interval; $\hat{\tau}$ = random-effects standard deviation estimate; M = unweighted mean; SD = unweighted standard deviation.

discrimination, regardless of how subtle, spontaneous, controlled, or deliberate they were.

Explicit measures of bias were also, on average, weak predictors of criteria in the studies covered by this meta-analysis, but explicit measures performed no worse than, and sometimes better than, the IATs for predictions of policy preferences, interpersonal behavior, person perceptions, reaction times, and microbehavior. Only for brain activity were correlations higher for IATs than for explicit measures ($\hat{\rho}$ = .42 vs. $\hat{\rho}$ = .34), but few studies examined prediction of brain activity using explicit measures. Any distinction between the IATs and explicit measures is a distinction that makes little difference, because both of these means of measuring attitudes resulted in poor prediction of racial and ethnic discrimination.

We do not consider the finding that the IAT correlated highly with brain activity criterion scores in the race domain to be evidence that the IAT is predictive of discriminatory behavior. In fact, we hesitated to include the neuroimaging studies in this meta-analysis, because this domain stands out as one in which null results are not reported in published studies and thus not available for meta-analytic review (see online supplemental materials) and because we cannot conceive of any socially meaningful definition of discrimination that treats differences in brain activity—independent of relevant behavioral outcomes—as discrimination (cf. Gazzaniga, 2005). We included these studies because Greenwald, Poehlman, et al. (2009) included brain scans as a proxy for discrimination and because we suspected, and found, that effects from fMRI studies would be large, thus inflating the aggregated IAT–criterion correlations relative to correlations for other criteria examined (smaller sample sizes notwithstanding). To be clear, neuroimaging studies of the biological origins of bias and differential processing of majority and minority targets may yield important theoretical and even practical insights, but without some empirical link to outward verbal or nonverbal behavior, these studies do not bear directly on the ability of the IAT to predict acts of racial and ethnic discrimination.

Flawed Theories or Flawed Instruments?

Why did the IAT and explicit measures perform so poorly with respect to all criteria that did not involve brain scan data? One explanation locates the problem in the instruments themselves, whereas an alternative explanation locates the problem in the theories that inspired the development and use of the instruments. We interpret our results as most consistent with the flawed instruments explanation, but our results also raise important questions for existing theories of implicit social cognition and prejudice.

Theoretical implications. The low predictive utility for the race and ethnicity IATs present problems for contemporary theories of prejudice and discrimination that assign a central role to implicit constructs (see Amodio & Mendoza, 2010). Explicitly endorsed ethnic and racial biases have become less common, yet societal inequalities persist. In response, psychologists have theorized that implicit biases must be a key sustainer of these inequalities (e.g., Chugh, 2004; Rudman, 2004), and IAT research has become the primary exhibit in support of this theory. The present results call for a substantial reconsideration of implicit-bias-based theories of discrimination at the level of operationalization and measurement, at least to the extent those theories depend on IAT research for proof of the prevalence of implicit prejudices and

Table 7
Meta-Analysis of ECCs by Criterion Scoring Method: Black Versus White Groups

Criterion scoring method	k (s; N_{total})	$\hat{\rho}$ [95% CI]	$\hat{\tau}$	M	SD
Overall					
Absolute—Black target$_a$	61 (19; 3,908)	.17 [.06, .29]	.17	.14	.19
Absolute—White target$_{a,b}$	35 (11; 1,520)	−.04 [−.12, .04]	.00	−.04	.16
Relative rating$_{b,c}$	17 (8; 3,791)	.17 [.10, .25]	.19	.16	.14
Difference score$_c$	97 (18; 4,487)	.06 [−.03, .14]	.10	.04	.18
Interpersonal behavior					
Absolute—Black target	8 (2; 664)	.31 [[a]]	.23	.32	.18
Absolute—White target	2 (1; 280)	−.10 [[a]]	—	−.10	.12
Relative rating	—	—	—	—	—
Difference score	2 (1; 280)	.01 [[a]]	—	.01	.02
Person perception					
Absolute—Black target$_a$	27 (9; 957)	.23 [.01, .45]	.21	.14	.17
Absolute—White target$_{a,b}$	25 (7; 925)	−.04 [−.15, .08]	.01	−.04	.18
Relative rating$_b$	10 (5; 542)	.19 [.08, .31]	.00	.16	.13
Difference score	17 (5; 1,021)	.04 [−.11, .19]	.12	.06	.17
Policy preference					
Absolute—Black target	18 (4; 1,966)	.08 [−.17, .34]	.17	.07	.18
Absolute—White target	—	—	—	—	—
Relative rating	3 (1; 3,171)	.25 [[a]]	—	.25	.15
Difference score	—	—	—	—	—
Microbehavior					
Absolute—Black target	7 (4; 300)	.13 [−.04, .29]	.07	.08	.18
Absolute—White target	7 (3; 295)	−.05 [−.17, .06]	.00	−.03	.13
Relative rating	4 (3; 78)	.09 [−.12, .30]	.00	.08	.15
Difference score	73 (8; 2,918)	.00 [−.12, .12]	.10	.01	.17
Response time					
Absolute—Black target	1 (1; 21)	.46 [[a]]	—	.46	—
Absolute—White target	1 (1; 20)	.19 [[a]]	—	.19	—
Relative rating	—	—	—	—	—
Difference score	3 (2; 243)	.18 [–.41, .77]	.00	.22	.13
Brain activity					
Absolute—Black target	—	—	—	—	—
Absolute—White target	—	—	—	—	—
Relative rating	—	—	—	—	—
Difference score	2 (2; 25)	.33 [[a]]	.20	.28	.33

Note. All effects were coded such that positive correlations are in the direction of promajority group or antiminority group responses or behaviors. The correlation between dependent effects is assumed to be .50. The $\hat{\rho}$ for each category is based on a moderated meta-analysis across categories, where dependent effect sizes (both within and across categories) are accounted for (Hedges et al., 2010), and the overall random-effects variance (tau-squared) weight is applied. $\hat{\tau}$ is also independently estimated within each category in separate analyses. Dashes indicate insufficient number of effects for computation purposes. Effects sharing subscripts within a category set are statistically significantly different from one another ($p < .05$). ECCs = explicit-criterion correlations; k = number of effects; s = number of independent samples within each category (this does not add up to the overall s because of sample overlap across categories); $\hat{\rho}$ = meta-analytically estimated population correlation; CI = confidence interval; $\hat{\tau}$ = random-effects standard deviation estimate; M = unweighted mean; SD = unweighted standard deviation.
[a] An appropriate estimate cannot be computed due to the integrated analysis with limited effects in this category.

assume that the prejudices measured by IATs are potent drivers of behavior. This conclusion follows as well from the broader meta-analytic results from Greenwald, Poehlman, et al. (2009), which reported weak evidence for implicit bias as a predictor of behavior in the gender and sexual orientation domain and which reported that the IAT explained small amounts of variance in an absolute sense within the domains of interracial and other intergroup behavior.

One might argue that, although the predictive utility of the IAT was low in most criterion domains, the existence of some weak, reliable effects might nonetheless be of interest to science if they advance basic theory (e.g., Mook, 1983; Prentice & Miller, 1992; Rosenthal & Rubin, 1979). However, it was not just the magnitude but also the pattern of effect sizes in the current analysis that are hard to reconcile with current theory. All theories of implicit social cognition, whether they embrace simple association or dissociation models of the relation of implicit constructs to behavior (Cameron et al., 2012; Perugini et al., 2010), hypothesize that implicitly measured constructs will, at a minimum, influence some spontaneous behaviors. Yet, the race and ethnicity IATs were weak or unreliable predictors of the more spontaneous behaviors covered by this meta-analysis. This finding raises questions about the proper conception of implicit bias (i.e., as more state-like or trait-like; cf. Smith & Conrey, 2007) and suggests that situational conditions can powerfully sway even the relationship between implicit bias and spontaneous behaviors. Across the board, the correlations observed between the IATs and criterion behaviors failed to reach the levels observed by Wicker (1969) in his classic

Table 8
Meta-Analysis of ECCs by Criterion Scoring Method: Ethnic Minority Versus Majority Groups

Criterion scoring method	k (s; N_{total})	$\hat{\rho}$ [95% CI]	$\hat{\tau}$	M	SD
Overall					
Absolute—Minority target	20 (9; 2,765)	.17 [.03, .30]	.14	.18	.19
Absolute—Majority target	3 (2; 102)	.04 [–.09, .16]	.00	.06	.09
Relative rating	—	—	—	—	—
Difference score	42 (6; 2,650)	.14 [–.05, .34]	.20	.11	.19
Interpersonal behavior					
Absolute—Minority target	1 (1; 105)	.18 [[a]]	—	.18	—
Absolute—Majority target	—	—	—	—	—
Relative rating	—	—	—	—	—
Difference score	—	—	—	—	—
Person perception					
Absolute—Minority target	8 (5; 212)	.04 [–.24, .31]	.19	.08	.26
Absolute—Majority target	3 (2; 102)	.04 [–.11, .18]	.00	.06	.09
Relative rating	—	—	—	—	—
Difference score	34 (4; 2,038)	.24 [–.05, .53]	.25	.14	.20
Policy preference					
Absolute—Minority target	10 (3; 2,343)	.23 [.16, .30]	.00	.22	.07
Absolute—Majority target	—	—	—	—	—
Relative rating	—	—	—	—	—
Difference score	—	—	—	—	—
Microbehavior					
Absolute—Minority target	1 (1; 105)	.47 [[a]]	—	.47	—
Absolute—Majority target	—	—	—	—	—
Relative rating	—	—	—	—	—
Difference score	8 (2; 612)	.01 [–.32, .33]	.00	.00	.07

Note. All effects were coded such that positive correlations are in the direction of promajority group or antiminority group responses or behaviors. The correlation between dependent effects is assumed to be .50. The $\hat{\rho}$ for each category is based on a moderated meta-analysis across categories, where dependent effect sizes (both within and across categories) are accounted for (Hedges et al., 2010), and the overall random-effects variance (tau-squared) weight is applied. $\hat{\tau}$ is also independently estimated within each category in separate analyses. Dashes indicate insufficient number of effects for computation purposes. No effects within a category set are statistically significantly different from one another ($p < .05$). No studies were available to examine the impact of criterion scoring method on correlations with response times or brain activity in the ethnicity domain. ECCs = explicit-criterion correlations; k = number of effects; s = number of independent samples within each category (this does not add up to the overall s because of sample overlap across categories); $\hat{\rho}$ = meta-analytically estimated population correlation; CI = confidence interval; $\hat{\tau}$ = random-effects standard deviation estimate; M = unweighted mean; SD = unweighted standard deviation.
[a] An appropriate estimate cannot be computed due to the integrated analysis with limited effects in this category.

review of attitude–behavior relations—levels that prompted soul-searching within social psychology about the attitude construct. The tremendous heterogeneity observed within and across criterion categories indicates that how implicit biases translate into behavior—if and when they do at all—appears to be complex and hard to predict.

Our findings, paired with Cameron et al.'s (2012) finding that sequential priming measures were better predictors of behavior in domains with higher correlations between implicit and explicit measures, suggest that the relation of implicit bias to behavior will be particularly weak in the domain of prejudice and discrimination. Cameron et al.'s results suggest that a lack of conflict between constructs accessed implicitly and explicitly translates into stronger behavioral effects. We did not observe such a pattern: In the ethnicity domain, where there was the highest correlation between measures, neither implicit nor explicit measures showed notably better prediction. Overall, implicit–explicit correlations were often quite low, with minuscule incremental validity. This result is not surprising, given that implicitly and explicitly measured intergroup attitudes so often diverge, and it suggests that one explanation for our results may be the existence of this implicit–explicit conflict. That is, the flip side of Cameron et al.'s finding may apply here. If true, this would indicate that implicitly measured intergroup biases are much less of a behavioral concern than many have worried—precisely because explicit attitudes often diverge from implicit attitudes. Such an oppositional process, in which explicit attitudes often win out in charged domains, is consistent with Petty and colleagues' metacognitive model of attitudes (e.g., Petty, Briñol, & DeMarree, 2007). This model posits that initial evaluations are checked by validity tags that develop over time, often through controlled processes such as conscious thought about one's views of a group, and the functioning of these tags can become automated over time and thus capable of checking even seemingly spontaneous behaviors.

One difficulty with this account for our findings—and with any theory that posits a role of conscious evaluations in the production of discrimination—is that even the explicit attitude measures in this domain offered weak prediction of meaningful criteria. In fact, one might argue that, given the much longer history of explicit than implicit attitude measurement, this meta-analysis strikes a sharper blow to traditional theories of prejudice by revealing the poor predictive utility of explicit attitude measures in the domains

Table 9
IAT (ICC) and Explicit Measures (ECC) Incremental Analysis: Percentage Variance Accounted for Across All Criteria

Explicit measure	ICC + ECC	ICC only	ECC only	ICC over ECC	ECC over ICC
Black vs. White					
Thermometer	3.2	2.3	1.2	2.0	0.9
Other existing	3.0	2.3	1.2	1.8	0.7
Created scale	2.4	2.3	0.4	2.0	0.1
Ethnic minority vs. majority groups					
Thermometer	1.6	1.4	0.4	1.2	0.2
Other existing	2.0	1.4	0.8	1.2	0.6
Created scale	6.8	1.4	5.8	1.0	5.4

Note. Analyses are based on relevant ICC, ECC, and IEC meta-analytic correlations reported in previous tables. ICC + ECC is the total $R^2 \times 100$; it is not a simple sum of their contributions to prediction, because it takes IECs into account. Results have been rounded to the nearest tenth of a percent. These analyses use only the difference score ICCs from Heider and Skowronski (2007) and Stanley et al. (2011). IAT = Implicit Association Test; ICC = implicit criterion-related validities (without any explicit measure above); ECC = explicit criterion-related validities (for the explicit measure listed); IEC = implicit–explicit correlation.

of ethnic and racial discrimination. Many of the studies examined here relied on published, theoretically-grounded measures of bias, including measures designed to assess modern racism (McConahay, 1986), symbolic racism (Henry & Sears, 2002), and ambivalent racism (Katz & Hass, 1988). Yet, these published, validated inventories fared no better than simple feeling thermometers or ad hoc instruments created by researchers. This is particularly worrisome from a theoretical point of view.

One potential reason why the more theoretically grounded instruments came up short is that they, like the IAT, seek to measure prejudicial attitudes indirectly and seek to capture subterranean racist motivations that can be hard to separate from nonracial motivations behind support for or opposition to various political policies (Sniderman & Tetlock, 1986). Perhaps the lack of ICC and ECC prediction argues for a reconsideration of this broader theoretical foundation. A more likely explanation, however, involves the use of these general attitude measures to predict a wide variety of criteria for which they were not compatible.

Instrument implications. Although our results suggest that amendments may be in order for theories of implicit social cognition and prejudice, a more parsimonious explanation lies with the instruments themselves. Our results give reason to believe that both the IATs and the explicit measures used in the criterion studies suffered from inherent limitations that compromised their criterion prediction, particularly the result that both types of measures failed to achieve validity levels comparable to those found by Wicker (1969) and in more recent meta-analyses of attitude–behavior relations (e.g., Kraus, 1995).

IAT measurement model. The IAT requires that two attitude objects be placed in opposition, as with Blacks and Whites on the race IAT. IAT researchers have argued that the relative nature of IAT measures can be a strength that enhances its predictive utility in certain criterion-prediction contexts (Nosek & Sriram, 2007). In contrast, we have shown that the difference-score nature of the IAT imposes a restrictive model that obscures the understanding and validity of its contributing components in most common criterion-prediction settings (Blanton, Jaccard, Christie, & Gonzales, 2007). The patterns observed here reinforce concerns introduced by Blanton et al. (2007) and call the dual-category format of the IAT into question in the domain of prejudice (cf. Pittinsky, 2010; Pittinsky et al., 2011). If the racial attitude IAT is a valid and reliable measure of the relative evaluations of Blacks compared to Whites, we should have found correlations of roughly equal magnitude between the race IAT and criterion measures, regardless of how criterion measures were scored and regardless of the race of the target (see Blanton, Jaccard, Gonzales, & Christie, 2006). Instead, for the interpersonal behavior and person perception criteria, the race IAT was a poor predictor of behavior toward Whites. ICCs were close to zero or negative for White-target-only criteria other than brain activity. Moreover, the high levels of moderate and strong bias as measured by scores on the IAT that are often found in the criterion studies, combined with the low levels of predictive validity found for all of the criterion behaviors, suggest that the dual-category approach does not measure attitude strength even for people whose associations with each object are strongest or the least in conflict (i.e., people whose difference scores on the IAT are greatest) or alternatively that the race and ethnicity IATs do not measure, at least not primarily, "attitudes" that have predictable psychological force and social meaning.

IAT metric. Nominally, bias on the IAT denotes only differential response times between the compatible and incompatible blocks. Most typically, IAT researchers score their measures so that positive scores are assumed to indicate bias against racial and ethnic minorities. The robust tendency—in most populations and measurement contexts—for most IAT measures of this type to yield more positive than negative scores has been broadly interpreted as evidence that implicit racial and ethnic biases are prevalent (e.g., Banaji & Greenwald, 2013). Despite the seeming face validity of such interpretations, researchers should not impute specific meaning to specific IAT response patterns prior to systematic empirical research that tests for potential links between different IAT scores and observable actions that can be understood in terms of the degree of racial or ethnic bias they reveal (Blanton & Jaccard, 2006). Without an independent means of validating current interpretations, it remains possible that the IAT is rank ordering individuals on one or more psychological constructs that can reliably reproduce positive scores across a wide range of populations and measurement contexts but do so for reasons having little to do with the modal distribution of implicit biases. Given evidence that the IAT in part measures skill at switching tasks, familiarity with different stimulus objects, and working memory capacity, and that it might be contaminated by other sources of method-based variance (e.g., Bluemke & Fiedler, 2009; Rothermund, Teige-Mocigemba, Gast, & Wentura, 2009; Teige-Mocigemba, Klauer, & Rothermund, 2008), many processes or constructs other than evaluative or semantic group-based associations may account for positive IAT scores. At the least, the low IAT–criterion correlations observed here counsel strongly against the assumption that scores on the race and ethnicity IATs reflect individual differences in propensity to discriminate.

Construction of explicit measures. The fact that the explicit measures were weak predictors of all criteria other than brain activity, and at levels below those found in prior meta-analyses of prejudice–behavior relations (Kraus, 1995; Talaska et al., 2008), supports the view that explicit measurement can be improved in IAT studies. Our review of the explicit bias measures used in the criterion studies leads us to echo statements made by Talaska et al. (2008) regarding the variable quality of the measures used in the criterion studies they synthesized. In particular, we agree with their concern about the lack of attention given to the compatibility between the component of prejudice being measured and the behavior being predicted. Given that explicit measures provide a standard against which the utility of implicit measures are evaluated and given that comparisons between the predictive utility of implicit and explicit measures are central to tests of theory, our analysis at a minimum points to the need for vigorous attention to improving the measurement of explicit attitudes in the IAT literature.

Criterion study implications. The low predictive validities we observed could also be due to limitations of the criterion studies apart from the limitations of the instruments. One potential limitation is restricted range on the criterion measures. We have observed low levels of discrimination across participants in some individual IAT criterion studies (Blanton et al., 2009; Blanton & Mitchell, 2011). It may be that many criterion studies in this meta-analysis contained little variance to be explained or predicted, which would prevent the discovery of high correlations. That would not vindicate the IAT's construct or predictive validity (because the high levels of bias implied by IAT scores led to many false-positive predictions of discrimination in the criterion studies), but it would hold out the possibility that the IAT could fare better in samples exhibiting more variability in levels of discrimination.

One possible source of restricted range is participants moderating their behavior to avoid appearing prejudiced. Researchers did not consistently report how their protocols might have masked the racial or ethnic implications of the tasks that respondents were asked to perform, and it is not unlikely that participants in many studies divined the purpose of the research. Some researchers did utilize unobtrusive observation of intergroup interactions as a main criterion (e.g., McConnell & Leibold, 2001), but for understandable practical reasons, these researchers often assessed implicit and explicit bias in the same experimental session as the criterion assessment, which may have sensitized participants to the general purpose of the study. Nevertheless, we discount this possibility as a general explanation for our results. The need to mask the purpose behind the study, to avoid reactivity bias and range restriction, points to something of a dilemma for researchers seeking to use explicit measures of bias that honor compatibility concerns: To the extent the measure taps into attitudes and beliefs more specific to the task and targets at hand, the more likely it is that participants will infer that the study seeks to examine prejudice and discrimination. This concern may partially explain the simple and general nature of many of the explicit measures used in the studies we synthesized, but one should be careful not to draw strong theoretical inferences about the relative strength of implicit and explicit measures from a measurement constraint imposed by laboratory settings and experimental requirements.

Another limitation of the criterion studies was their consistently small sample sizes. Most of the individual-study effects were based on sample sizes below 50; the median sample sizes for the overall ICC, ECC, and IEC were 41, 41, and 77, respectively. These small sample sizes yield correlation estimates that have large associated margins of error ("MOE," which equals half the width of the 95% confidence interval), making most individual correlation estimates imprecise. For example, given a correlation of 0.20, the MOEs for sample sizes of the level typically found in IAT research are as follows: for N of 25, the MOE is ± 0.38 correlation units; for $N = 50$, MOE $= \pm 0.27$; for $N = 75$, MOE $= \pm 0.22$; for $N = 100$, MOE $= \pm 0.19$. In the rare case where $N \geq 250$, then MOE $\leq \pm 0.12$ (note that these are average MOEs, because they are asymmetric due to use of the Fisher r-to-z transformation). To provide more acceptable MOEs in correlational studies, one should use sample sizes of at least $N = 250$.

Future Directions

There are many steps researchers should consider taking not only to improve prediction but also to deepen their understanding of prejudice-behavior relations. For instance, it may be possible to improve the predictive validity of the IAT by examining more closely method-specific variance. The IAT's designers favor an algorithm-based approach to artifacts that seeks to minimize the influence of known confounds post hoc through the use of scoring algorithms (Greenwald, Nosek, & Banaji, 2003). A more cumbersome but perhaps more effective strategy would be to measure and statistically control for known confounds. Researchers should consider developing portfolios of independent measures that assess the influence of systematic confounds, so that their influences can be statistically assessed and either modeled as covariates or substantively controlled in future research designs.

We also recommend that future comparative research on explicit and implicit attitudes adopt latent variable modeling that can accommodate multiple measures of the same construct (whether implicit or explicit), so that results are less measure dependent and thus less confounded with inferences about constructs and their relationships (e.g., Nosek & Smyth, 2007). Such models can accommodate measurement-error variance due to random errors of measurement and the idiosyncratic features of any given measure; they can also accommodate transient error found in longitudinal analysis as well as other forms of error-variance structure. Even the most generous estimates of the reliability of IAT variants consistently show them to be lower than those for explicit measures (Bosson, Swann, & Pennebaker, 2000; Cunningham, Preacher, & Banaji, 2001). Researchers should therefore bring methods to bear that can more effectively separate measurement error from the attitudinal signal or true score of interest.

Greater consideration should also be given to reducing demand characteristics and other sources of reactivity bias, by designing experimental protocols that are appropriately neutral and by inserting time lags and distractor tasks between measures of intergroup attitudes and behaviors. Survey researchers have outlined procedures for minimizing social desirability bias, including (a) ensuring respondents of the confidentiality and privacy of their responses, (b) allowing respondents to complete questions under conditions of anonymity, (c) stressing the importance of honest and candid responses, (d) using honesty pledges as part of the informed consent, (e) avoiding face-to-face reporting of answers to socially sensitive questions, (f) obtaining measures of social desirability as either trait or response tendencies that can be examined as correlates or covariates during statistical modeling, and (g) building in experimental tests for differential response patterns (see Tourangeau & Yan, 2007).

Study documentation could be improved in several ways, too. As with all psychology studies, researchers should report the results for all their criterion measures and scoring procedures, rather than aggregate the criteria and report results for only a single scoring method (Fiedler, 2011). Building a stronger cumulative body of knowledge about the relation of implicit and explicit bias to discriminatory behavior will require disentangling these dynamics (i.e., documenting effects of implicit and explicit attitudes on pro-White behavior, on anti-Black behavior, and on Black–White behavior differentials), rather than merging them through global behavior reports at the study level and meta-analytic effects averaged across heterogeneous outcomes. As a conceptual matter, relative or comparative criterion measures of some kind should always be included in studies of discriminatory behavior in order to determine whether differential treatment has indeed occurred. Criterion studies that focus on treatment of a minority target may provide important data on interpersonal relations, but such studies do not support inferences of racial or ethnic discrimination (see Blanton & Mitchell, 2011, and Blanton et al., 2009, for examples of interpretive problems that arise from not reporting comparative results).

Finally, future meta-analyses of social psychological studies of phenomena may benefit from the meta-analytic approach that we adopted, in which multiple effects from a single sample can be included by taking into account dependencies among these effects. In many social psychology studies, the focus is on how behavior changes across situations or tasks. An approach in which a single effect is calculated from averaging across within-sample effects is not necessary, and it causes the loss of important information about substantive variation across effects.

Conclusion

The initial excitement over IAT effects gave rise to a hope that the IAT would prove to be a window on unconscious sources of discriminatory behavior. This hope has been sustained by individual studies finding statistically significant correlations between IAT scores and some criterion measures of discrimination and by the finding from Greenwald, Poehlman, et al. (2009) that IATs had greater predictive validity than explicit measures of bias when predicting discrimination against African Americans and other minorities. This closer look at the IAT criterion studies in the domains of ethnic and racial discrimination revealed, however, that the IAT provides little insight into who will discriminate against whom, and provides no more insight than explicit measures of bias. The IAT is an innovative contribution to the multidecade quest for subtle indicators of prejudice, but the results of the present meta-analysis indicate that social psychology's long search for an unobtrusive measure of prejudice that reliably predicts discrimination must continue (see Crosby, Bromley, & Saxe, 1980; Mitchell & Tetlock, in press). Overall, simple explicit measures of bias yielded predictions no worse than the IATs. Had researchers attended to the compatibility principle in the development of the explicit measures and consistently taken steps to minimize reactivity bias, the explicit measures would likely have performed substantially better (cf. Kraus, 1995; Talaska et al., 2008).

References

References marked with an asterisk are included in the meta-analysis.

Agerström, J., & Rooth, D. (2011). The role of automatic obesity stereotypes in real hiring discrimination. *Journal of Applied Psychology, 96,* 790–805. doi:10.1037/a0021594

*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. doi:10.1037/0022-3514.91.4.652

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

Arkes, H., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or "Would Jesse Jackson fail the Implicit Association Test?" *Psychological Inquiry, 15,* 257–278. doi:10.1207/s15327965pli1504_01

*Ashburn-Nardo, L., Knowles, M. L., & Monteith, M. J. (2003). Black Americans' implicit racial associations and their implications for intergroup judgment. *Social Cognition, 21,* 61–87. doi:10.1521/soco.21.1.61.21192

*Avenanti, A., Sirigu, A., & Aglioti, S. M. (2010). Racial bias reduces empathic sensorimotor resonance with other-race pain. *Current Biology, 20,* 1018–1022. doi:10.1016/j.cub.2010.03.071

Banaji, M. R. (2008, August). The science of satire: Cognition studies clash with "New Yorker" rationale. *Chronicle of Higher Education, 54,* B13.

Banaji, M. R., & Greenwald, A. G. (2013). *Blindspot: Hidden biases of good people*. New York, NY: Random House.

Bar-Anan, Y., & Nosek, B. A. (2012). *A comparative investigation of seven implicit measures of social cognition*. Unpublished manuscript, University of Virginia.

Bennett, M. W. (2010). Unraveling the Gordian knot of implicit bias in jury selection: The problems of judge-dominated voir dire, the failed promise of Batson, and proposed solutions. *Harvard Law & Policy Review, 4,* 149–171.

*Biernat, M., Collins, E. C., Katzarska-Miller, I., & Thompson, E. R. (2009). Race-based shifting standards and racial discrimination. *Personality and Social Psychology Bulletin, 35,* 16–28. doi:10.1177/0146167208325195

Biggerstaff, B. J., & Tweedie, R. L. (1997). Incorporating variability in estimates of heterogeneity in the random effects model in meta-analysis. *Statistics in Medicine, 16,* 753–768. doi:10.1002/(SICI)1097-0258(19970415)16:7<753::AID-SIM494>3.0.CO;2-G

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist, 61,* 27–41. doi:10.1037/0003-066X.61.1.27

Blanton, H., Jaccard, J., Christie, C., & Gonzales, P. M. (2007). Plausible assumptions, questionable assumptions, and post hoc rationalizations: Will the real IAT please stand up? *Journal of Experimental Social Psychology, 43,* 399–409. doi:10.1016/j.jesp.2006.10.019

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. doi:10.1016/j.jesp.2005.07.003

Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., & Tetlock, P. E. (2009). Strong claims and weak evidence: Reassessing the predictive validity of the IAT. *Journal of Applied Psychology, 94,* 567–582. doi:10.1037/a0014665

Blanton, H., & Mitchell, G. (2011). Reassessing the predictive validity of the IAT II: Reanalysis of Heider & Skowronski (2007). *North American Journal of Psychology, 13,* 99–106.

Blasi, G., & Jost, J. T. (2006). System justification theory and research: Implications for law, legal advocacy, and social justice. *California Law Review, 94,* 1119–1168. doi:10.2307/20439060

Bluemke, M., & Fiedler, K. (2009). Base rate effects on the IAT. *Consciousness and Cognition, 18,* 1029–1038. doi:10.1016/j.concog.2009.07.010

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, England: Wiley. doi:10.1002/9780470743386

Bosson, J. K., Swann, W. B., Jr., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79,* 631–643. doi:10.1037/0022-3514.79.4.631

Cameron, C. D., Brown-Iannuzzi, J., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: A meta-analysis of associations with behaviors and explicit attitudes. *Personality and Social Psychology Review, 16,* 330–350. doi:10.1177/1088868312440047

*Carney, D. R. (2006). *The faces of prejudice: On the malleability of the attitude–behavior link*. Unpublished manuscript, Harvard University.

*Carney, D. R., Olson, K. R., Banaji, M. R., & Mendes, W. B. (2006). *The faces of race-bias: Awareness of racial cues moderates the relation between bias and in-group facial mimicry*. Unpublished manuscript, Harvard University.

Cheung, S. F., & Chan, D. K.-S. (2004). Dependent effect sizes in meta-analysis: Incorporating the degree of interdependence. *Journal of Applied Psychology, 89,* 780–791. doi:10.1037/0021-9010.89.5.780

Cheung, S. F., & Chan, D. K.-S. (2008). Dependent correlations in meta-analysis: The case of heterogeneous dependence. *Educational and Psychological Measurement, 68,* 760–777. doi:10.1177/0013164408315263

Chugh, D. (2004). Societal and managerial implications of implicit social cognition: Why milliseconds matter. *Social Justice Research, 17,* 203–222. doi:10.1023/B:SORE.0000027410.26010.40

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. doi:10.1037/0022-3514.83.6.1314

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302. doi:10.1037/h0040957

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. doi:10.1037/0033-2909.87.3.546

*Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of Black and White faces. *Psychological Science, 15,* 806–813. doi: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. doi:10.1111/1467-9280.00328

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

Ditonto, T. M., Lau, R. R., & Sears, D. O. (in press). AMPing racial attitudes: Comparing the power of explicit and implicit racism measures in 2008. *Political Psychology*.

Dovidio, J. F., Kawakami, K., Smoak, N., & Gaertner, S. L. (2009). The nature of contemporary racial prejudice: Insight from implicit and explicit measures of attitudes. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 165–192). New York, NY: Psychology Press.

Drummond, M. A. (Spring, 2011). ABA Section of Litigation tackles implicit bias. *Litigation News, 36,* 20–21.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75–109). New York, NY: Academic Press. doi:10.1016/S0065-2601(08)60318-4

Fiedler, K. (2011). Voodoo correlations are everywhere—Not only in social neurosciences. *Perspectives on Psychological Science, 6,* 163–171. doi:10.1177/1745691611400237

Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. New York, NY: Psychology Press.

*Florack, A., Scarabis, M., & Bless, H. (2001). Der Einfluß wahrgenommener Bedrohung auf die Nutzung automatischer Assoziationen bei der Personenbeurteilung. *Zeitschrift für Sozialpsychologie, 32,* 249–259. doi:10.1024//0044-3514.32.4.249

Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, H., & Hart, A. (2004) A threat in the computer: The race Implicit Association Test as a stereotype threat experience. *Personality and Social Psychology Bulletin, 30,* 1611–1624. doi:10.1177/0146167204266650

*Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: Associations influence the construal of individuating information. *European Journal of Social Psychology, 33,* 573–589. doi:10.1002/ejsp.166

Gazzaniga, M. S. (2005). *The ethical brain*. New York, NY: Dana Press.

Gladwell, M. (2005). *Blink: The power of thinking without thinking*. New York, NY: Little, Brown.

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

Gleser, L. J., & Olkin, I. (2007). *Stochastically dependent effect sizes* (Department of Statistics, Technical Report No. 2007–2). Stanford, CA: Stanford University.

*Green, A. R., Carney, D. R., Pallin, D. J., Ngo, L. H., Raymond, K. L., Iezzoni, L. I., & Banaji, M. R. (2007). Implicit bias among physicians and its prediction of thrombolysis decisions for Black and White patients. *Journal of General Internal Medicine, 22,* 1231–1238. doi:10.1007/s11606-007-0258-5

Greenwald, A. G. (2006, September 3). *Expert report of Anthony G. Greenwald, Satchell v. FedEx Express,* No. C 03–2659 (N. D. Cal.).

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

Greenwald, A. G., & Krieger, L. H. (2006). Implicit bias: Scientific foundations. *California Law Review, 94,* 945–967. doi:10.2307/20439056

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. doi: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. doi:10.1037/0022-3514.85.2.197

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,* 17–41. doi:10.1037/a0015575

*Greenwald, A. G., Smith, C. T., Sriram, N., Bar-Anan, Y., & Nosek, B. A. (2009). Implicit race attitudes predicted vote in the 2008 U.S. presidential election. *Analyses of Social Issues and Public Policy, 9,* 241–253. doi:10.1111/j.1530-2415.2009.01195.x

Hartung, J., & Knapp, G. (2003). An alternative test procedure for meta-analysis. In R. Schultze, H. Holling, & D. Böhning (Eds.), *Meta-analysis: New developments and applications in medical and social sciences* (pp. 53–69). Cambridge, MA: Hofgrefe & Huber.

*He, Y., Johnson, M. K., Dovidio, J. F., & McCarthy, G. (2009). The relation of race-related implicit associations and scalp-recorded neural activity evoked by faces from different races. *Social Neuroscience, 4,* 426–442. doi:10.1080/17470910902949184

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1,* 39–65. doi:10.1002/jrsm.5

*Heider, J. D., & Skowronski, J. J. (2007). Improving the predictive validity of the Implicit Association Test. *North American Journal of Psychology, 9,* 53–76.

Henry, P. J., & Sears, D. O. (2002). The Symbolic Racism 2000 Scale. *Political Psychology, 23,* 253–283. doi:10.1111/0162-895X.00281

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. doi:10.1177/0146167205275613

*Hofmann, W., Gschwendner, T., Castelli, L., & Schmitt, M. (2008). Implicit and explicit attitudes and interracial interaction: The moderating role of situationally available control resources. *Group Processes & Intergroup Relations, 11,* 69–87. doi:10.1177/1368430207084847

*Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science, 14,* 640–643. doi:10.1046/j.0956-7976.2003.psci_1478.x

*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. doi:10.1111/j.0956-7976.2004.00680.x

*Hughes, J. M., & Bigler, R. S. (2011). Predictors of African American and European American adolescents' endorsement of race-conscious social policies. *Developmental Psychology, 47,* 479–492. doi:10.1037/a0021309

Irwin, J. F., & Real, D. L. (2010). Unconscious influences on judicial decision-making: The illusion of objectivity. *McGeorge Law Review, 42,* 1–18.

Jaccard, J., & Blanton, H. (2006). A theory of implicit reasoned action: The role of implicit and explicit attitudes in the prediction of behavior. In I. Ajzen, D. Albarracin, & J. Hornik (Eds.), *Prediction and change of health behavior: Applying the reasoned action approach* (pp. 53–68). Mahwah, NJ: Erlbaum.

Kang, J. (2005). Trojan horses of race. *Harvard Law Review, 118,* 1489–1593.

Kang, J., Bennett, M. W., Carbado, D. W., Casey, P., Dasgupta, N., Faigman, D., . . . Mnookin, J. (2012). Implicit bias in the courtroom. *UCLA Law Review, 59,* 1124–1186.

*Kang, J., Dasgupta, N., Yogeeswaran, K., & Blasi, G. (2010). Are ideal litigators White? Measuring the myth of colorblindness. *Journal of Empirical Legal Studies, 7,* 886–915. doi:10.1111/j.1740-1461.2010.01199.x

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81,* 774–788. doi:10.1037/0022-3514.81.5.774

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structure. *Journal of Personality and Social Psychology, 55,* 893–905. doi:10.1037/0022-3514.55.6.893

Kim, R.-S., & Becker, B. J. (2010). The degree of dependence between multiple-treatment effect sizes. *Multivariate Behavioral Research, 45,* 213–238. doi:10.1080/00273171003680104

*Korn, H., Johnson, M. A., & Chun, M. M. (2012). Neurolaw: Differential brain activity for Black and White faces predicts damage awards in hypothetical employment discrimination cases. *Social Neuroscience, 7,* 398–409. doi:10.1080/17470919.2011.631739

Kraus, S. J. (1995). Attitudes and the prediction of behavior: A meta-analysis of the empirical literature. *Personality and Social Psychology Bulletin, 21,* 58–75. doi:10.1177/0146167295211007

*Levinson, J. D., Cai, H., & Young, D. M. (2010). Guilty by implicit bias: The guilty/not guilty implicit association test. *Ohio State Journal of Criminal Law, 8,* 187–208.

Levinson, J. D., & Smith, R. J. (Eds.). (2012). *Implicit racial bias across the law*. Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511820595

Levinson, J. D., Young, D. M., & Rudman, L. A. (2012). Implicit racial bias: A social science overview. In J. D. Levinson & R. J. Smith (Eds.), *Implicit racial bias across the law* (pp. 9–24). Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511820595.002

*Livingston, R. W. (2002). *Bias in the absence of malice: The paradox of unintentional deliberative discrimination*. Unpublished manuscript, University of Wisconsin—Madison.

*Ma-Kellams, C., Spencer-Rodgers, J., & Peng, K. (2011). I am against us? Unpacking cultural differences in ingroup favoritism via dialecticism. *Personality and Social Psychology Bulletin, 37,* 15–27. doi:10.1177/0146167210388193

*Maner, J. K., Kenrick, D. T., Becker, D. V., Robertson, T. E., Hofer, B., Neuberg, S. L., . . . Schaller, M. (2005). Functional projection: How fundamental social motives can bias interpersonal perception. *Journal of Personality and Social Psychology, 88,* 63–78. doi:10.1037/0022-3514.88.1.63

McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–125). San Diego, CA: Academic Press.

*McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37,* 435–442. doi:10.1006/jesp.2000.1470

Mitchell, G., & Tetlock, P. E. (2006). Antidiscrimination law and the perils of mindreading. *Ohio State Law Journal, 67,* 1023–1121.

Mitchell, G., & Tetlock, P. E. (in press). Implicit attitude measures. In R. A. Scott & S. M. Kosslyn (Eds.), *Emerging trends in the social and behavioral sciences*. Thousand Oaks, CA: Sage.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38,* 379–387. doi:10.1037/0003-066X.38.4.379

Nosek, B. A., & Smyth, F. L. (2007). A multitrait–multimethod validation of the Implicit Association Test: Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology, 54,* 14–29. doi:10.1027/1618-3169.54.1.14

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., . . . Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology, 18,* 36–88. doi:10.1080/10463280701489053

Nosek, B. A., & Sriram, N. (2007). Faulty assumptions: A comment on Blanton, Jaccard, Gonzales, and Christie (2006). *Journal of Experimental Social Psychology, 43,* 393–398. doi:10.1016/j.jesp.2006.10.018

Olson, M. A., & Fazio, R. H. (2009). Implicit and explicit measures of attitudes: The perspective of the MODE model. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 19–63). New York, NY: Psychology Press.

Oswald, F. L., & Johnson, J. W. (1998). On the robustness, bias, and stability of results from meta-analysis of correlation coefficients: Some initial Monte Carlo findings. *Journal of Applied Psychology, 83,* 164–178. doi:10.1037/0021-9010.83.2.164

Page, A., & Pitts, M. J. (2009). Poll workers, election administration, and the problem of implicit bias. *Michigan Journal of Race & Law, 15,* 1–56. doi:10.2139/ssrn.1392630

Parks, G. S., & Rachlinski, J. J. (2010). Implicit bias, election '08, and the myth of a post-racial America. *Florida State University Law Review, 37,* 659–715.

*Pérez, E. O. (2010). Explicit evidence on the import of implicit attitudes: The IAT and immigration policy judgments. *Political Behavior, 32,* 517–545. doi:10.1007/s11109-010-9115-z

*Perugini, M., O'Gorman, R., & Prestwich, A. (2007). An ontological test of the IAT: Self-activation can increase predictive validity. *Experimental Psychology, 54,* 134–147. doi:10.1027/1618-3169.54.2.134

Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In B. Gawronski & K. Payne (Eds.), *Handbook of implicit social cognition* (pp. 255–277). New York, NY: Guilford Press.

Petty, R. E., Briñol, P., & DeMarree, K. G. (2007). The meta-cognitive model (MCM) of attitudes: Implications for attitude measurement, change, and strength. *Social Cognition, 25,* 657–686. doi:10.1521/soco.2007.25.5.657

*Phelps, E. A., O'Conner, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12,* 729–738. doi:10.1162/089892900562552

Pittinsky, T. L. (2010). A two-dimensional model of intergroup leadership: The case of national diversity. *American Psychologist, 65,* 194–200. doi:10.1037/a0017329

Pittinsky, T. L., Rosenthal, S. A., & Montoya, R. M. (2011). Liking is not the opposite of disliking: The functional separability of positive and negative attitudes toward minority groups. *Cultural Diversity and Ethnic Minority Psychology, 17,* 134–143. doi:10.1037/a0023806

Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. *Psychological Bulletin, 112,* 160–164. doi:10.1037/0033-2909.112.1.160

*Prestwich, A., Kenworthy, J. B., Wilson, M., & Kwan-Tat, N. (2008). Differential relations between two types of contact and implicit and explicit racial attitudes. *British Journal of Social Psychology, 47,* 575–588. doi:10.1348/014466607X267470

Quillian, L. (2006). New approaches to understanding racial prejudice and discrimination. *Annual Review of Sociology, 32,* 299–328. doi:10.1146/annurev.soc.32.061604.123132

*Rachlinski, J. J., Johnson, S. L., Wistrich, A. J., & Guthrie, C. (2009). Does unconscious racial bias affect trial judges? *Notre Dame Law Review, 84,* 1195–1246.

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

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

Rosenthal, R., & Rubin, D. B. (1979). A note on percent of variance explained as a measure of the importance of effects. *Journal of Applied Social Psychology, 9,* 395–396. doi:10.1111/j.1559-1816.1979.tb02713.x

Rothermund, K., Teige-Mocigemba, S., Gast, A., & Wentura, D. (2009). Minimizing the influence of recoding in the Implicit Association Test: The Recoding-Free Implicit Association Test (IAT-RF), *Quarterly Journal of Experimental Psychology, 62,* 84–98. doi:10.1080/17470210701822975

Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test: Dissociating salience from associations. *Journal of Experimental Psychology: General, 133,* 139–165. doi:10.1037/0096-3445.133.2.139

Rudman, L. A. (2004). Social justice in our minds, homes, and society: The nature, causes, and consequences of implicit bias. *Social Justice Research, 17,* 129–142. doi:10.1023/B:SORE.0000027406.32604.f6

*Rudman, L. A., & Ashmore, R. D. (2007). Discrimination and the Implicit Association Test. *Group Processes & Intergroup Relations, 10,* 359–372. doi:10.1177/1368430207078696

*Rudman, L. A., & Lee, M. R. (2002). Implicit and explicit consequences of exposure to violent and misogynous rap music. *Group Processes & Intergroup Relations, 5,* 133–150. doi:10.1177/1368430202005002541

*Sabin, J. A., Rivara, F. P., & Greenwald, A. G. (2008). Physician implicit attitudes and stereotypes about race and quality of medical care. *Medical Care, 46,* 678–685. doi:10.1097/MLR.0b013e3181653d58

*Sargent, M. J., & Theil, A. (2001). *When do implicit racial attitudes predict behavior? On the moderating role of attributional ambiguity*. Unpublished manuscript, Bates College.

Scheck, J. (2004, October 28). *Expert witness: Bill Bielby helped launch an industry–suing employers for unconscious bias*. Retrieved from http://www.law.com/jsp/PubArticle.jsp?id=900005417471

Schnabel, K., Asendorpf, J. B., & Greenwald, A. G. (2008). Understanding and using the Implicit Association Test: V. Measuring semantic aspects of trait self-concepts. *European Journal of Personality, 22,* 695–706. doi:10.1002/per.697

Sears, D. O. (2004a). Continuities and contrasts in American racial politics. In J. T. Jost, M. R. Banaji, & D. A. Prentice (Eds.), *Perspectivism in social psychology: The yin and yang of scientific progress* (pp. 233–245). Washington, DC: APA Books. doi:10.1037/10750-017

Sears, D. O. (2004b). A perspective on implicit prejudice from survey research. *Psychological Inquiry, 15,* 293–297.

Sears, D. O., & Henry, P. J. (2005). Over thirty years later: A contemporary look at symbolic racism and its critics. *Advances in Experimental Social Psychology, 37,* 95–150. doi:10.1016/S0065-2601(05)37002-X

*Sekaquaptewa, D., Espinoza, P., Thompson, M., Vargas, P., & von Hippel, W. (2003). Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination. *Journal of Experimental Social Psychology, 39,* 75–82. doi:10.1016/S0022-1031(02)00512-7

*Shelton, J. N., Richeson, J. A., Salvatore, J., & Trawalter, S. (2005). Ironic effects of racial bias during interracial interactions. *Psychological Science, 16,* 397–402.

Shermer, M. (2006, November 24). Comic's outburst reflects humanity's sin. *Los Angeles Times*. Retrieved from Westlaw Newsroom, 2006 WLNR 20385662.

Shin, P. S. (2010). Liability for unconscious discrimination? A thought experiment in the theory of employment discrimination law. *Hastings Law Journal, 62,* 67–101.

*Sibley, C. G., Liu, J. H., & Khan, S. S. (2010). Implicit representations of ethnicity and nationhood in New Zealand: A function of symbolic or resource-specific policy attitudes? *Analyses of Social Issues and Public Policy, 10,* 23–46. doi:10.1111/j.1530-2415.2009.01197.x

Smith, E. R., & Conrey, F. R. (2007). Mental representations are states, not things: Implications for implicit and explicit measurement. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 247–264). New York, NY: Guilford Press.

Sniderman, P. M., & Tetlock, P. E. (1986). Symbolic racism: Problems of motive attribution in political debate. *Journal of Social Issues, 42,* 129–150. doi:10.1111/j.1540-4560.1986.tb00229.x

*Spicer, C. V., & Monteith, M. J. (2001). *Implicit outgroup favoritism among African Americans and vulnerability to stereotype threat*. Unpublished manuscript.

*Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences, USA, 108,* 7710–7715. doi:10.1073/pnas.1014345108

*Stepanikova, I., Triplett, J., & Simpson, B. (2011). Implicit racial bias and prosocial behavior. *Social Science Research, 40,* 1186–1195. doi:10.1016/j.ssresearch.2011.02.004

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8,* 220–247. doi:10.1207/s15327957pspr0803_1

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. B., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: Implications for clinical practice. *American Psychologist, 62,* 271–286. doi:10.1037/0003-066X.62.4.271

Talaska, C. A., Fiske, S. T., & Chaiken, S. (2008). Legitimating racial discrimination: A meta-analysis of the racial attitude–behavior literature shows that emotions, not beliefs, best predict discrimination. *Social Justice Research, 21,* 263–296. doi:10.1007/s11211-008-0071-2

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2008). Minimizing method-specific variance in the IAT: A single block IAT. *European Journal of Psychological Assessment, 24,* 237–245. doi:10.1027/1015-5759.24.4.237

Tetlock, P. E., & Mitchell, G. (2009). Implicit bias and accountability systems: What must organizations do to prevent discrimination? *Research in Organizational Behavior, 29,* 3–38. doi:10.1016/j.riob.2009.10.002

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133,* 859–883. doi:10.1037/0033-2909.133.5.859

*Tuttle, K. M. K. (2009). *Implicit racial attitudes and law enforcement shooting decisions*. Unpublished manuscript, University of Michigan.

*Vanman, E. J., Saltz, J. L., Nathan, L. R., & Warren, J. A. (2004). Racial discrimination by low-prejudiced Whites: Facial movements as implicit measures of attitudes related to behavior. *Psychological Science, 15,* 711–714. doi:10.1111/j.0956-7976.2004.00746.x

Vedantam, S. (2010). *The hidden brain: How our unconscious minds elect presidents, control markets, wage wars, and save our lives*. New York, NY: Random House.

*Vezzali, L., & Giovannini, D. (2011). Intergroup contact and reduction of explicit and implicit prejudice toward immigrants: A study with Italian businessmen owning small and medium enterprises. *Quality & Quantity: International Journal of Methodology, 45,* 213–222. doi:10.1007/s11135-010-9366-0

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science, 4,* 274–290. doi:10.1111/j.1745-6924.2009.01125.x

Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues, 25,* 41–78. doi:10.1111/j.1540-4560.1969.tb00619.x

*Ziegert, J. C., & Hanges, P. J. (2005). Employment discrimination: The role of implicit attitudes, motivation, and a climate for racial bias. *Journal of Applied Psychology, 90,* 553–562. doi:10.1037/0021-9010.90.3.553

Received August 26, 2011
Revision received March 18, 2013
Accepted March 18, 2013 ■