Personality and Social Psychology Bulletin

http://psp.sagepub.com/

Attitudinal and Non-Attitudinal Components of IAT Performance

Jimmy Calanchini, Jeffrey W. Sherman, Karl Christoph Klauer and Calvin K. Lai

Pers Soc Psychol Bull 2014 40: 1285 originally published online 1 July 2014
DOI: 10.1177/0146167214540723

The online version of this article can be found at:
http://psp.sagepub.com/content/40/10/1285

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

On behalf of:

Society for Personality and Social Psychology

Article

Personality and Social
Psychology Bulletin
2014, Vol. 40(10) 1285–1296

DOI: 10.1177/0146167214540723
pspb.sagepub.com
SAGE

Attitudinal and Non-Attitudinal Components of IAT Performance

Jimmy Calanchini[1], Jeffrey W. Sherman[1], Karl Christoph Klauer[2], and Calvin K. Lai[3]

Abstract

The Implicit Association Test (IAT) was designed to measure automatically activated attitudinal associations, free of the influence of processes that affect their expression. Subsequent research has shown that IAT performance also is influenced by non-associative processes, but the extent to which these non-associative processes are content-specific or if they operate similarly regardless of the attitude being measured has largely gone unexamined. In the current research, participants completed pairs of IATs that varied in conceptual overlap: Tests shared a high, moderate, or low degree of overlap in the measured attitudes. The Quad model was applied to estimate the contributions of four processes to IAT performance. Evidence was found for two relatively general, non-attitudinal processes and two relatively attitude-specific processes. Implications are discussed for interpretation of IAT scores, individual differences in IAT scores, and IAT score malleability.

Keywords

implicit cognition, stereotyping, prejudice, quantitative models

Received April 8, 2013; revision accepted May 30, 2014

The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) was designed to measure automatically activated attitudinal associations, free of the influence of processes that affect the expression of those associations. Specifically, because it is less susceptible to self-presentational demands and is more able to capture attitudes that are inaccessible through introspection than are self-report measures (Gawronski, 2009), the IAT has been viewed as a more "pure" measure of attitudinal associations that are stored in long-term memory. Indeed, the IAT has been successful in predicting a host of behavioral, judgmental, and physiological outcomes (for a review, see Greenwald, Poehlman, Uhlmann, & Banaji, 2009), and has been applied to myriad research domains, including prejudice (e.g., Gawronski, Peters, Brochu, & Strack, 2008), stereotyping (e.g., Gawronski, Ehrenberg, Banse, Zukova, & Klauer, 2003), self-esteem (e.g., Greenwald & Farnham, 2000), self-concept (e.g., Asendorpf, Banse, & Mücke, 2002), brand evaluation (e.g., Forehand & Perkins, 2005), mental health (e.g., Teachman, Gregg, & Woody, 2001), and addiction (e.g., Wiers, Houben, & de Kraker, 2007; Wiers, Van Woerden, Smulders, & De Jong, 2002), to name just a few.

However, subsequent research has shown that IAT performance is not influenced solely by activated attitudinal associations; a wide variety of non-associative processes also influence IAT performance. Some of these processes include the inhibition of associations, the detection of correct responses, response biases (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Sherman et al., 2008), recoding processes (e.g., Chang & Mitchell, 2011; De Houwer, Geldof, & De Bruycker, 2005; Kinoshita & Peek-O'Leary, 2005, 2006; Meissner & Rothermund, 2013; Rothermund, Teige-Mocigemba, Gast, & Wentura, 2009; Rothermund & Wentura, 2001, 2004; Rothermund, Wentura, & De Houwer, 2005), task-set shifts and task-set simplification (e.g., Klauer, Schmitz, Teige-Mocigemba, & Voss, 2010; Mierke & Klauer, 2001, 2003), and speed-accuracy tradeoffs (e.g., Brendl, Markman, & Messner, 2001; Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007).

An important unanswered question is the extent to which these non-associative processes are attitude-specific (i.e., they operate differently depending on the attitude being measured) or are content-general (i.e., they operate similarly regardless of the specific attitude being measured). In the former case, a non-associative process may still be regarded as construct-related (i.e., attitude-specific), in that its operation would be

[1]University of California, Davis, USA
[2]University of Freiburg, Germany
[3]University of Virginia, Charlottesville, USA

Corresponding Author:
Jimmy Calanchini, University of California, Davis, 134 Young Hall, 1 Shields Ave., Davis, CA 95616, USA.
Email: jcalanchini@ucdavis.edu

tied to a specific attitudinal object. However, in the latter case, a process would instead be unrelated to the nature of the attitude object, and would reflect general cognitive processes that influence task performance independently of the construct being measured. We refer to such a content-generic process as "non-attitudinal" in that it would reflect method variance rather than construct variance.

The degree to which IAT performance is driven by attitude-specific versus non-attitudinal processes has important implications for understanding IAT effects, individual differences in IAT scores, the susceptibility of IAT scores to change, and the ability of IAT scores to predict important behaviors. For example, Gonsalkorale, Sherman, and Klauer (2009) presented data suggesting that older people show greater racial IAT bias than younger people not because of more strongly biased associations, but because of decreased regulation of their biased associations. The interpretation of these findings depends greatly on whether the diminished regulation of biased associations among older respondents is specific to the domain of race (and is, therefore, attitude-specific) or is a general deficit that applies across attitude objects (and is, therefore, non-attitudinal). If the diminished regulation is not attitude-specific, then Gonsalkorale and colleagues' findings would seem to reflect age-related deficits in general cognitive function that have nothing to do with racial bias, per se (Connelly, Hasher, & Zacks, 1991; Hasher & Zacks, 1988). However, if the deficit is specific to the domain of race, then these findings suggest other interpretations entirely. For example, older people may be unwilling, rather than unable, to constrain their biased responses, or perhaps simply grew up in an era in which they were never taught to constrain racial bias and, therefore, lack the ability to regulate only in this domain.

Similarly, the degree to which the processes that drive IAT performance are general versus attitude-specific has implications for understanding the effectiveness and generalizability of bias-reduction interventions. For example, an intervention that targets a process that is attitude-specific should not be expected to effectively reduce bias in other content domains. However, an intervention that targets a process that is domain-general should be expected to reduce bias in other content domains. Thus, it is important to understand the specificity versus generality of the processes that underlie implicit bias to guide the development of bias-reduction interventions.

The question of attitude-specificity versus domain-generality is also important for conceptualizing and measuring attitudes, more broadly. If the non-associative processes that influence IAT performance are attitude-specific, then those processes may be viewed as important components of the evaluative response to that attitude object. In contrast, if these non-associative processes are non-attitudinal, then this would suggest that they are irrelevant to the evaluative response to the attitude object and that their influence should be removed in measuring those evaluative responses.

We are aware of very little research that directly investigates the attitude-specific versus non-attitudinal nature of the processes that contribute to evaluative responses on implicit measures of attitude (but see Klauer et al., 2010; Mierke & Klauer, 2001, 2003). The purpose of the current research was to directly examine this question in the context of IAT performance. The Quadruple Process model (Quad model; Sherman et al., 2008) is especially well suited for such a process-level investigation.

The Quad Model

According to the Quad model, four distinct processes influence performance on implicit tasks like the IAT (i.e., response-conflict tasks): Activation of Associations (AC), Detection of correct responses (D), Overcoming Bias (OB), and Guessing (G). The AC parameter refers to the degree to which biased associations are activated when responding to a stimulus. All else equal, the stronger the associations, the more likely they are to be activated and to drive behavior in an association-consistent direction. The D parameter reflects the likelihood that the respondent can discriminate between correct and incorrect responses. Sometimes, the activated associations conflict with the detected correct response. For example, on incompatible trials of an IAT (e.g., when the categories "Black" and "pleasant" share a response key), activated racial associations (e.g., between Black and unpleasant) conflict with the detected correct response (i.e., to press the same button for Black and pleasant stimuli). In such cases, the Quad model proposes that an OB process resolves the conflict. As such, the OB parameter refers to a regulatory process that prevents activated associations from influencing behavior when they conflict with detected correct responses. Finally, the G parameter reflects general response tendencies that may occur when individuals have no associations that direct behavior and they are unable to detect the correct response. The Quad model and the construct validity of its parameters have been extensively validated in previous research and have been shown to play important roles in IAT performance (see Beer et al., 2008; Conrey et al., 2005; Gonsalkorale, von Hippel, Sherman, & Klauer, 2009; Sherman et al., 2008).

Predictions

Though the Quad model parameters have been extensively validated, the extent to which these processes are attitude-specific versus domain-general remains an open question. We expect that AC is an attitude-specific process, as AC represents the likelihood that associations are activated by the specific targets presented in a given IAT (Sherman et al., 2008). For example, if flowers are not strongly associated with African Americans, there should be no strong relationship between an AC parameter that represents associations activated by flowers and an AC parameter that represents associations activated by African Americans.

D, OB, and G reflect non-associative processes that operate during IAT performance and help determine the extent to which responses are based on underlying associations. Though these processes are not associative in nature, they may or may not operate differently depending on the target construct that is being considered. D represents the likelihood of determining correct responses in a task such as the IAT (Sherman et al., 2008). It is possible that the ability to detect a correct response represents a general motivation or cognitive skill that does not vary as a function of target construct. If so, D would be a non-attitudinal process, and D in one content domain should be related to D in another domain. Alternatively, the ability to detect a correct response may represent attitude-specific abilities such as domain expertise. For example, a florist may be especially skilled at discerning correct responses on a flower/insect IAT, but may perform at an average level on a race IAT. If this is the case, then D would be a relatively attitude-specific process, and D in one domain should be unrelated to D in other domains.

OB represents the likelihood that an activated association is overcome and replaced by a contextually accurate response (Sherman et al., 2008). Like D, it is possible that OB is either an attitude-specific or a domain-general process. OB is conceptualized as an inhibitory process that constrains activated associations from driving responses. Because inhibition is often conceptualized as a domain-general ability (e.g., prepotent response inhibition; Friedman & Miyake, 2004), this suggests that OB would be a non-attitudinal process, and the ability to overcome bias in one content domain should be related to that ability in other domains. Alternatively, it is possible that the likelihood of overcoming biased associations varies across domains based on a variety of social, cultural, or personal pressures. For example, one might be motivated to regulate biased responses toward sensitive topics (i.e., racial/ethnic relations) but not toward non-sensitive topics (e.g., insects), which could affect the extent of OB observed in the two domains. In this case, OB would be a relatively attitude-specific process, and OB in one domain should be unrelated to OB in other domains.

G represents biases to prefer a particular response, such as a positivity bias (e.g., Conrey et al., 2005) or a preference for objects on the right side of a display (e.g., Nisbett & Wilson, 1977). It is unclear whether such a bias should be specific to certain domains or span across domains. For example, G might generalize into an optimistic outlook on everything when conceived as a positivity bias (e.g., "I'm just a positive person"). In this case, G would be a non-attitudinal process, and G in one content domain should be related to G in other domains. However, G may vary across content domains as even a positive outlook might have limits (e.g., "I like people but I hate spiders."). In this case, G in one domain should be unrelated to G in other domains, and would be considered an attitude-specific process.

Overview of Research

In order to test the attitude-specificity versus generality of the Quad parameters, participants completed pairs of IATs that varied in conceptual overlap. High conceptual overlap was operationalized as IATs with the same attribute categories (e.g., pleasant, unpleasant) and target categories from the same content domain (e.g., race). Moderate conceptual overlap was operationalized as IATs with the same attribute categories but target categories from different content domains (e.g., disability, sexual orientation). Low conceptual overlap was operationalized as IATs with different attribute categories that also had target categories from different content domains.

By examining the relations among parameters estimated from IATs of varying conceptual overlap, the attitude-specific or general nature of a given process can be discerned. A domain-general, non-attitudinal process should be insensitive to conceptual overlap and, therefore, should be expected to have correlation coefficients that are large in magnitude across IATs, regardless of conceptual overlap. Conversely, an attitude-specific process should be sensitive to conceptual overlap, and should be expected to have larger correlation coefficients across IATs with high conceptual overlap than IATs with low conceptual overlap. Of course, any process may reflect varying degrees of both attitude-specificity and generality.

The current research is structured as an exploratory/confirmatory design. We entered into this program of research with clear *a priori* predictions about the attitude-specificity of AC, but with competing predictions for D, OB, and G. Thus, we sought initial evidence for one set of predictions over the other in Exploratory Studies 1a, 1b, and 1c, and then attempted to replicate these findings in Confirmatory Studies 2a, 2b, and 2c. Importantly, the exploratory and confirmatory studies were conducted on different participant samples and used different instantiations of conceptual overlap.

Study 1

Overview and Procedure

The purpose of Study 1 was to explore the attitude-specificity versus generality of the Quad processes by varying the conceptual overlap of pairs of IATs. Participants completed pairs of IATs that either shared high (Study 1a), moderate (Study 1b), or low (Study 1c) conceptual overlap. All studies were completed in a laboratory setting at the University of California, Davis, and participants received either partial course credit or US$10 payment.

The structure of all of the IATs in Study 1 was based on that described by Greenwald et al. (1998). Participants first completed a 20-trial practice block in which they categorized attribute words (e.g., pleasant, unpleasant) followed by a 20-trial practice block in which they categorized target stimuli (e.g., images of Black or White people). In the third

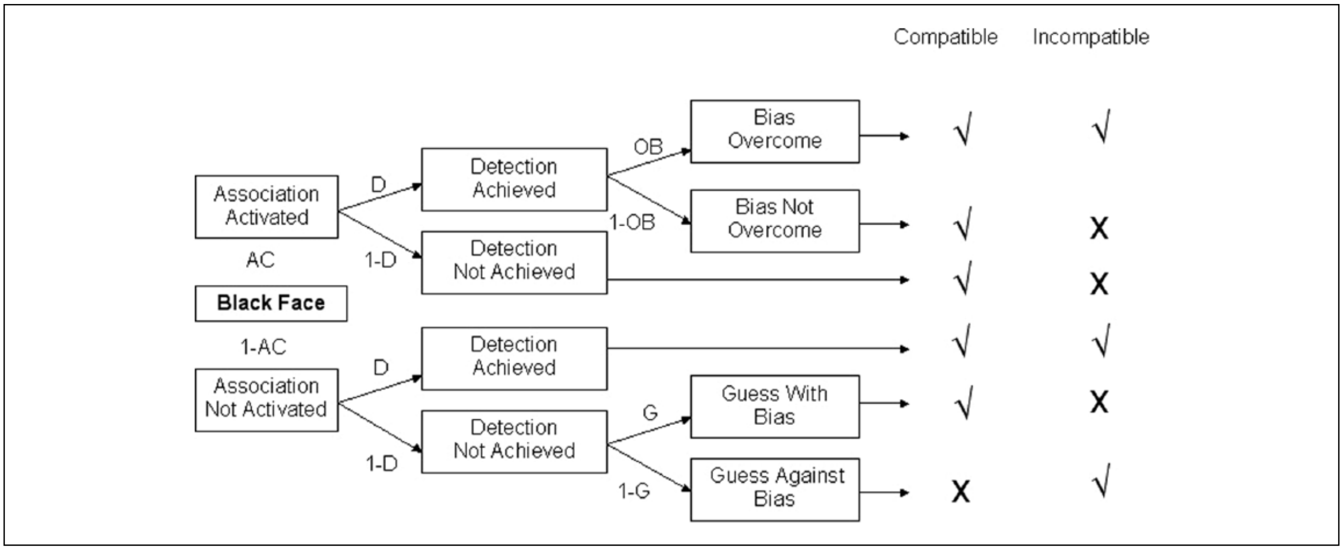

Figure 1. The Quadruple Process model (Quad model).
Note. Each path represents a likelihood. Parameters with lines leading to them are conditional upon all preceding parameters. The table on the right side of the figure depicts correct (√) and incorrect (X) responses as a function of process pattern.

(24 trial) and fourth (48 trial) critical blocks, participants simultaneously categorized both attribute and target stimuli. The fifth block was identical to the second block, but the response keys for the target stimuli were reversed. The sixth and seventh critical blocks were identical to the third and fourth blocks, but the response keys reflected the reversed target pairing of the fifth block. If a participant made an error in categorization during any of the response trials, a red "X" appeared below the stimulus and remained there until the participant corrected the error. The order in which the IATs were administered was counterbalanced between participants (except for Study 1c), and participants were given a 3-min break between IATs.

Study 1a

Participants and materials. Participants were 168 undergraduates (74% female) who completed two IATs that shared high conceptual overlap: a Black/White evaluative IAT and an Asian/White evaluative IAT. The Black/White evaluative IAT consisted of pleasant and unpleasant words and images of Black and White males. The target category labels were Black/White. The Asian/White evaluative IAT consisted of pleasant and unpleasant words and images of Asian and White females. The target category labels were Asian/White. These IATs not only share an attribute category (pleasant/unpleasant) but also measure attitudes toward distinct but related targets (i.e., they both measure intergroup bias).

Results. The Quad model has been implemented as a multinomial model (see Batchelder & Riefer, 1999) designed to estimate the independent contributions of multiple processes from responses on implicit measures of bias (for reviews of this approach, see Sherman, 2006; Sherman et al., 2008). The structure of the Quad model is depicted as a processing tree in Figure 1. In the tree, each path represents a likelihood. Processing parameters with lines leading to them are conditional on all preceding parameters. For instance, OB is conditional on both AC and D. The conditional relationships described by the model form a system of equations that predicts the numbers of correct and incorrect responses in different conditions (e.g., compatible and incompatible trials). For example, there are three ways in which an incorrect response can be returned on an incompatible trial, in which "Black" and "pleasant" share a response key. The first is the likelihood that biased associations are activated (AC), detection succeeds (D), and OB fails (1 − OB), which can be represented by the equation $AC \times D \times (1 - OB)$. The second is the likelihood that the biased associations are activated (AC) and detection fails (1 − D), which can be represented by the equation $AC \times (1 - D)$. The third is the likelihood that biased associations are not activated (1 − AC), detection fails (1 − D), and a bias toward guessing "unpleasant" (1 − G) produces an incorrect response, which can be represented by the equation $(1 - AC) \times (1 - D) \times (1 - G)$. As such, the overall likelihood of producing an incorrect response on an incompatible trial is the sum of these three conditional probabilities: $[AC \times D \times (1 - OB)] + [AC \times (1 - D)] + [(1 - AC) \times (1 - D) \times (1 - G)]$. The respective equations for each item category (e.g., White, Black, pleasant words, and unpleasant words in both compatible and incompatible blocks) are then used to predict the observed proportions of errors in a given data set. The model's predictions are compared to the actual data to determine the model's ability to account for the data. A chi-square estimate is computed for the difference between the predicted and observed errors. To best approximate the model to the data, the parameter values are changed through

Table 1. Cross-Test Correlation Coefficients.

	Average AC	D	OB	G
1a: High overlap	0.58 [0.31, 0.77]	0.82 [0.72, 0.89]	0.76 [0.34, 0.93]	0.36 [−0.21, 0.68]
1b: Moderate overlap	0.48 [0.08, 0.81]	0.71 [0.45, 0.87]	−0.20 [−0.76, 0.74]	0.02 [−0.68, 0.64]
1c: Low overlap	0.36 [0.06, 0.64]	0.68 [0.33, 0.88]	0.43 [−0.24, 0.87]	0.02 [−0.65, 0.72]

Note. 95% credible intervals in brackets. AC = Activation of Associations; D = Detection of correct responses; OB = Overcoming Bias; G = Guessing.

maximum likelihood estimation until they produce a minimum possible value of the chi-square. The final parameter values that result from this process are interpreted as relative levels of the processes.

Parameter estimates of AC, D, OB, and G were calculated for each participant for each test. The G parameter was coded so that higher scores represented a bias toward guessing with the "pleasant" key. Four separate AC parameters were estimated. Using responses from the Black/White evaluative IAT, one parameter estimated the extent to which associations between "Black" and "unpleasant" were activated and another estimated the extent to which associations between "White" and "pleasant" were activated. Using responses from the Asian/White evaluative IAT, one parameter estimated the extent to which associations between "Asian" and "unpleasant" were activated and another estimated the extent to which associations between "White" and "pleasant" were activated. The overall error rate for the Black/White IAT was 8.35%, and the chi-square for model fit was 16.36, $p = .001$. The overall error rate for the Asian/White IAT was 7.25%, and the chi-square for model fit was 24.73, $p < .001$. Chi-square tests are dependent on sample size, such that minute deviations from the model can jeopardize model fit when power is high (see Cohen, 1988). However, the effect size of lack of model fit between the actual data and the model's predicted data was small for the Black/White IAT, $w = .03$, and for the Asian/White IAT, $w = .03$, indicating satisfactory fit when controlling for power.

Comparing parameters across tests is, unfortunately, not as straightforward as calculating Pearson correlation coefficients. The different model parameters stem from the same data for each IAT and, consequently, their error terms are not independent. This may lead to artifactual dependencies between the parameters of each IAT, and failure to control for them may bias the assessment of relationships between the two different IATs. In addition, the individual parameter estimates are based on relatively few data points per person, implying that their reliability will be relatively low, limiting the ability of individual-level analyses to detect any relationships there may be.

Both these problems can be solved by the hierarchical multinomial model approach proposed by Klauer (2010). In this analysis, the model parameters are defined as latent traits, as in a structural equation model, and the correlations between the parameters of the two IATs as path coefficients, as in a structural equation model. In this analysis, dependencies between model parameters due to non-independent errors are controlled for, and the estimated correlations are automatically disattenuated for measurement error (again, as in analyses using structural equation models). The analysis uses a Bayesian approach, which yields estimates of both the correlations and the credible intervals (CIs) around them, which can be interpreted much like the familiar confidence intervals from classical statistics.

AC White-pleasant from the Black/White IAT correlated with AC White-pleasant from the Asian/White IAT, $r = .60$, $p < .001$, 95% CI [.31, .93]; AC White-pleasant from the Black/White IAT correlated with AC Asian-unpleasant, $r = .58$, $p < .001$, 95% CI [.14, .83]; AC Black-unpleasant correlated with AC White-pleasant from the Asian/White IAT, $r = .59$, 95% CI [.31, .80]; AC Black-unpleasant correlated with AC Asian-unpleasant, $r = .56$, 95% CI [.12, .81]. The average between-test AC correlation was $r = .58$, 95% CI [.31, .77]. Black/White D correlated with Asian/White D, $r = .82$, 95% CI [.72, .89]. Black/White OB correlated with Asian/White OB, $r = .76$, 95% CI [.34, .93]. Black/White G did not reliably correlate with Asian/White G, $r = .36$, 95% CI [−.21, .68]. These results are summarized in the top portion of Table 1.

Study 1b

Participants and materials. Participants were 49 undergraduates (71% female) who completed two IATs that shared moderate conceptual overlap: a flower/insect evaluative IAT and a Black/White evaluative IAT. The Black/White evaluative IAT was identical to the one described in Study 1a. The flower/insect evaluative IAT consisted of pleasant and unpleasant words and images of flowers and insects. The target category labels were flower/insect. These IATs share an attribute category (pleasant/unpleasant) but do not share a target category.

Results. Parameter estimates of AC, D, OB, and G were calculated for each participant for each test. One D, one OB, and one G parameter were estimated for each test. The G parameter was coded so that higher scores represented a bias toward guessing with the "pleasant" key. Four separate AC parameters were estimated. Using responses from the flower/insect evaluative IAT, one parameter estimated the extent to which associations between "insect" and "unpleasant" were activated and another estimated the extent to which associations between "flower" and "pleasant" were activated. Using responses from the Black/White evaluative IAT, one parameter estimated the extent to which associations between "Black" and "unpleasant" were activated and another estimated the extent to which associations between "White" and "pleasant" were activated. The overall error rate for the flower/insect IAT was 10.3%, and the chi-square for model fit was 5.47, $p = .14$, $w = .03$. The overall error rate for the racial IAT was 7.5%, and the chi-square for model fit was 6.30, $p = .10$, $w = 03$.

AC Black-unpleasant correlated with AC flower-pleasant, $r = .69$, 95% CI [.28, .91]; AC White-pleasant did not reliably correlate with AC flower-pleasant, $r = .47$, 95% CI [−.41, .88]; AC Black-unpleasant did not reliably correlate with AC insect-unpleasant, $r = .58$, 95% CI [−.17, .89]; AC White-pleasant did not reliably correlate with AC insect-unpleasant, $r = .35$, 95% CI [−.50, .83]. The average between-test AC correlation was $r = .48$, 95% CI [.08, .81]. Black/White D correlated with flower/insect D, $r = .71$, 95% CI [.45, .87]. Black/White OB did not reliably correlate with flower/insect OB, $r = −.20$, 95% CI [−.76, .74]. Black/White G did not reliably correlate with flower/insect G, $r = .02$, 95% CI [−.68, .64]. These results are summarized in the middle portion of Table 1.

Study 1c

Participants and materials. Participants were 56 undergraduates (61% female) who completed two IATs that shared low conceptual overlap: a Black/White stereotype IAT and a flower/insect evaluative IAT. The flower/insect evaluative IAT was identical to the one used in Study 1b. The Black/White stereotype IAT was based on one described by Amodio and Devine (2006). As in a Black/White evaluative IAT, a Black/White stereotype IAT uses images of Black and White faces as stimuli. However, in contrast to a Black/White evaluative IAT, the critical trials on a Black/White stereotype IAT pair the faces with words associated with Black and White stereotypes. African Americans are stereotyped as more athletic and less intelligent than European Americans (Devine & Elliot, 1995). Importantly, the categories *physical* and *mental* are evaluatively equivalent; therefore, categorization according to these traits does not involve evaluative judgments (see Amodio & Devine, 2006, footnote 3). Participants paired images of Black and White faces with physical and mental words that were taken from Amodio and Devine (2006, Study 1). The target category labels were Black/White and the attribute category labels were Physical/Mental. Participants always completed the Black/White stereotype IAT before the flower/insect evaluative IAT. These IATs did not share either attribute categories or target categories.

Results. Parameter estimates of AC, D, OB, and G were calculated for each participant for each test. The G parameter was coded so that higher scores represented a bias toward guessing with the "mental" key on the Black/White stereotype IAT and the "pleasant" key on the flower/insect evaluative IAT. Thus, in contrast to the previous studies, the G parameter represents qualitatively different processes on the different IATs. Four separate AC parameters were estimated. Using responses from the Black/White stereotype IAT, one AC parameter estimated the extent to which associations between "Black" and "physical" were activated in performing the task and another estimated the extent to which associations between "White" and "mental" were activated. Using responses from the flower/insect evaluative IAT, one AC parameter estimated the extent to which associations between "insect" and "unpleasant" were activated and another estimated the extent to which associations between "flower" and "pleasant" were activated. The overall error rate for the racial stereotype IAT was 8.77%, and the chi-square for model fit was 20.06, $p < .001$, $w = .05$. The overall error rate for the flower/insect evaluative IAT was 8.93%, and the chi-square for model fit was 1.45, $p = .69$, $w = .01$.

AC White-mental correlated with AC flower-pleasant, $r = .55$, 95% CI [.07, .81]; AC White-mental did not reliably correlate with AC insect-unpleasant, $r = .23$, 95% CI [−.21, .62]; AC Black-physical did not reliably correlate with AC flower-pleasant, $r = .45$, 95% CI [−.15, .87]; AC Black-physical did not reliably correlate with AC insect-unpleasant, $r = .25$, 95% CI [−.25, .69]. The average between-test AC correlation was $r = .36$, 95% CI [.06, .64]. Racial stereotype D correlated with flower/insect evaluative D, $r = .68$, 95% CI [.33, .88]. Racial stereotype OB did not reliably correlate with flower/insect evaluative OB, $r = .43$, 95% CI [−.24, .87]. Racial stereotype G did not reliably correlate with flower/insect evaluative G, $r = .02$, 95% CI [−.65, .72]. These results are summarized in the bottom portion of Table 1.

Summary and Discussion

The purpose of Study 1 was to explore the attitude-specificity versus generality of the Quad model processes by comparing the correlations among parameters derived from IATs with varying conceptual overlap. The specificity or generality of a given process can be assessed by examining the magnitude of these correlations. A general, non-attitudinal process should be insensitive to conceptual overlap, and should be expected to have correlation coefficients that are large in magnitude across IATs, regardless of conceptual overlap. Conversely, an attitude-specific process should be sensitive to conceptual overlap, and should be expected to have larger correlation coefficients across IATs with high conceptual overlap than IATs with low conceptual overlap. From this perspective, D appears to be a relatively general, non-attitudinal process, with correlations ranging from .68 to .80. Based on the guidelines recommended by Cohen (1992), this represents a large effect. Surprisingly, the pattern of correlations in Study 1a to 1c suggests that AC is also a relatively non-attitudinal process, with correlations ranging from .23 to .60. However, the magnitudes of the average AC correlations are smaller than the magnitude of the D correlations, and range from medium to large effects, which perhaps suggests a degree of attitude-specificity in AC as well. Less clear patterns of results emerged for the OB and G parameters. The relatively high magnitude of the OB correlations in the high and low overlap conditions (.76 and .43, respectively) suggests a degree of domain-generality, but the negative correlation in the moderate overlap condition (−.20) is inconsistent with this characterization. The moderate magnitude of the G correlation in the high overlap condition (.36) and near-zero correlations in the moderate and low overlap conditions (.02 in each case) suggest a degree of domain specificity. However, most of the OB correlations and all of the G correlations were not different from 0, which may reflect a very high degree of domain specificity in each of these processes, or it may, instead, simply reflect low reliability. In comparison to the D and AC parameters, OB and G are estimated from fewer IAT trials and are, therefore, inherently less reliable.

We found mixed evidence in exploratory Study 1 for our prediction that AC is an attitude-specific process. Regarding our competing predictions for D, OB, and G, we found relatively straightforward evidence that D is non-attitudinal, and mixed evidence for the attitude-specificity or generality of OB and G. In Study 2, we sought to further clarify the nature of these processes using different versions of the IAT and increased statistical power.

Study 2

Overview and Procedure

Moving into the confirmatory phase of the current research, we have two main goals for Study 2. The first goal is to replicate the pattern of domain-generality that emerged for the D parameter. The second is to clarify the pattern of results for the AC, OB, and G parameters. Study 2 is structured like Study 1, in that participants completed IATs of varying conceptual overlap. However, high, medium, and low overlap was instantiated in Study 2 using different attitude objects than were used in Study 1. Additionally, Study 2 was conducted on a much larger and more diverse participant sample. Because these confirmatory studies were conducted on a different sample using different instantiations of conceptual overlap, it is unlikely that the same set of biasing factors would distort the results in both sets of studies in the same directions. Thus, any similarities in the pattern of results obtained in the exploratory and confirmatory phases should be due to variations in conceptual overlap and, thus, reflect variations in domain-generality and -specificity among the Quad model parameters.

Participants were visitors to the Project Implicit website (https://implicit.harvard.edu) from 2006 to 2010 who chose to complete two of IATs from a list of up to 15 possible IATs. Each session lasted about 10 min and also contained questionnaires about attitudes, stereotypes, and judgments related to the topic of the IAT. The order of the questionnaires was randomized in most sessions, as was the order of the critical blocks of the IAT. Participants received feedback about their IAT performance at the end of each session.

The structure of the IATs was based on that described by Nosek, Banaji, and Greenwald (2002). Participants first completed a 20-trial practice block in which they categorized target stimuli followed by a 20-trial practice block in which they categorized attribute words. In the third (20 trial) and fourth (40 trial) critical blocks, participants simultaneously categorized both attribute and target stimuli. The fifth 40-trial block consisted of categorizing target stimuli, but with the response keys reversed. The sixth and seventh critical blocks were identical to the third and fourth blocks, but the response keys reflected the switched target pairing of the fifth block. If a participant made an error in categorization during any of the response trials, a red "X" appeared below the stimulus and remained there until the participant corrected the error.

Study 2a

Participants and materials. Participants were 33,278 visitors (52% female, 29% did not report) who completed two IATs that shared high conceptual overlap: a Black/White evaluative IAT and a skin tone evaluative IAT. The Black/White evaluative IAT consisted of pleasant and unpleasant words and images of Black and White males. The target images were different from the images used in Study 1a. The target category labels were either Black people/White people, African American/European American, Black American/White American, or Black/White. The skin tone evaluative IAT consisted of pleasant and unpleasant words and images of racially ambiguous light-skinned and dark-skinned people. The category labels for the target groups were Light Skin/Dark Skin. These IATs not only share an attribute category, but they also measure attitudes toward distinct but closely related social categories (Maddox, 2004).

Results. Parameter estimates of AC, D, OB, and G were calculated for each participant for each test. The G parameter was coded so that higher scores represented a bias toward guessing with the "pleasant" key. Four separate AC parameters were estimated. Using responses from the Black/White IAT, one parameter estimated the extent to which associations between "Black" and "unpleasant" were activated and another estimated the extent to which associations between "White" and "pleasant" were activated. Using responses from the skin tone IAT, one parameter estimated the extent to which associations between "dark skin" and "unpleasant" were activated and another estimated the extent to which associations between "light skin" and "pleasant" were activated. The overall error rate for the Black/White IAT was 8.35%, and the chi-square for model fit was 3,984.05, $p < .001$, $w = .03$. The overall error rate for the skin tone IAT was 8.84%, and the chi-square for model fit was 4,123.38, $p < .001$, $w = .03$.

Table 2. Cross-Test Correlation Coefficients.

	Average AC	D	OB	G
2a: High overlap	0.50 [0.49, 0.52]	0.67 [0.66, 0.68]	0.67 [0.59, 0.68]	0.21 [0.14, 0.26]
2b: Moderate overlap	0.37 [0.33, 0.41]	0.73 [0.71, 0.74]	0.85 [0.79, 0.92]	0.05 [−0.04, 0.11]
2c: Low overlap	0.20 [0.16, 0.24]	0.78 [0.77, 0.80]	0.66 [0.58, 0.89]	0.08 [−0.03, 0.16]

Note. 95% credible intervals in brackets. AC = Activation of Associations; D = Detection of correct responses; OB = Overcoming Bias; G = Guessing.

Because the Project Implicit samples are so large, power is virtually 100% and, consequently, even trivial effects are statistically reliable. As such, we focus on the magnitude of the effects, using Cohen's (1992) recommendations as guidelines. All four of the AC correlations were large effects: AC White-pleasant/AC light skin-pleasant, $r = .54$, 95% CI [.53, .56]; AC White-pleasant/AC dark skin-unpleasant, $r = .49$, 95% CI [.47, .51]; AC Black-unpleasant/AC light skin-pleasant, $r = .47$, 95% CI [.45, .50]; AC Black-unpleasant/AC dark skin-unpleasant, $r = .50$, 95% CI = [.48, .53]. The average between-test AC correlation was $r = .50$, 95% CI = [.49, .52], a large effect. The correlation between Black/White D and skin tone D was also a large effect, $r = .67$, 95% CI [.66, .68]. The correlation between Black/White OB and skin tone OB was a large effect, $r = .67$, 95% CI [.59, .68]. The correlation between Black/White G and skin tone G was a small-to-medium effect, $r = .21$, 95% CI [.14, .26]. These results are summarized in the top portion of Table 2.

Study 2b

Participants and materials. Participants were 9,570 visitors (62% female, 18% did not report) who completed two IATs that shared moderate conceptual overlap: a gay/straight evaluative IAT and disability evaluative IAT. The gay/straight IAT consisted of pleasant and unpleasant words and images representing gay and straight people. The target category labels were Straight People/Gay People. The disability evaluative IAT consisted of pleasant and unpleasant words and images representing abled or disabled people. The target category labels were Abled Persons/Disabled Persons. The gay/straight evaluative IAT used different good/bad words than the disability evaluative IAT. These IATs share an attribute category but measure attitudes toward unrelated target categories.

Results. Parameter estimates of AC, D, OB, and G were calculated for each participant for each test. One D, one OB, and one G parameter were estimated for each test. The G parameter was coded so that higher scores represented a bias toward guessing with the "pleasant" key. Four separate AC parameters were estimated. Using responses from the gay/straight IAT, one parameter estimated the extent to which associations between "gay" and "unpleasant" were activated and another estimated the extent to which associations between "straight" and "pleasant" were activated. Using responses from the disability IAT, one parameter estimated the extent to which associations between "disabled" and "unpleasant" were activated and another estimated the extent to which associations between "abled" and "pleasant" were activated. The overall error rate for the gay/straight IAT was 11.3%, and the chi-square for model fit was 2,465.03, $p < .001$, $w = .05$. The overall error rate for the disability IAT was 9.7%, and the chi-square for model fit was 2,925.55, $p < .001$, $w = .05$.

All four of the AC correlations were medium-to-large effects: AC straight-pleasant/AC abled-pleasant, $r = .36$, 95% CI [.30, .41]; AC straight-pleasant/AC disabled-unpleasant, $r = .34$, 95% CI [.27, .40]; AC gay-unpleasant/AC abled-pleasant, $r = .35$, 95% CI [.29, .40]; AC gay-unpleasant/AC disabled-unpleasant $r = .43$, 95% CI [.38, .48]. The average between-test AC correlation was $r = .37$, 95% CI [.33, .44], a medium-to-large effect. The correlation between gay/straight D and disability D was a large effect, $r = .73$, 95% CI [.71, .74]. The correlation between Black/White OB and skin tone OB was a large effect, $r = .85$, 95% CI [.79, .92]. The correlation between Black/White G and skin tone G was not different from 0, $r = .05$, 95% CI [−.04, .11]. These results are summarized in the middle portion of Table 2.

Study 2c

Participants and materials. Participants were 9,820 visitors (65% female, 16% did not report) who completed two IATs that shared low conceptual overlap: an age evaluative IAT and a gender-career stereotype IAT. The age evaluative IAT consisted of pleasant and unpleasant words and images of young and old people. The target category labels were Young/Old. The gender-career stereotype IAT consisted of words representing family and career (e.g., wedding, corporation) and male and female names (e.g., John, Michelle). The attribute category labels were Family/Career and the target category labels were Male/Female. These IATs did not share either attribute categories or target categories.

Results. Parameter estimates of AC, D, OB, and G were calculated for each participant for each test. The G parameter was coded so that higher scores represented a bias toward guessing with the "career" key on the gender-career IAT and the "pleasant" key on the age evaluative IAT. Thus, like Study 1c, the G parameter represents qualitatively different processes on the different IATs. Four separate AC parameters were estimated. Using responses from the gender-career IAT, one AC parameter estimated the extent to which associations between "female" and "family" were activated in performing the task and another estimated the extent to which associations between "male" and "career" were activated. Using responses from the age IAT, one AC parameter estimated the extent to which associations between "young" and "pleasant" were activated and another estimated the extent to which associations between "old" and "unpleasant" were activated. The overall error rate for the gender-career IAT was 8.85%, and the chi-square for model fit was 1,989.72, $p < .001$, $w = .04$. The overall error rate for the age evaluative IAT was 8.39%, and the chi-square for model fit was 815.31, $p < .001$, $w = .03$.

All four of the AC correlations were small-to-medium effects: AC male-career/AC young-pleasant, $r = .18$, 95% CI [.13, .25]; AC male-career/AC old-unpleasant, $r = .12$, 95% CI [.04, .18]; AC female-family/AC young-pleasant, $r = .31$, 95% CI [.25, .36]; AC female-family/AC old-unpleasant $r = .20$, 95% CI [.15, .26]. The average between-test AC correlation was $r = .20$, 95% CI [.16, .24], a small-to-medium effect. The correlation between gender-career D and age D was a large effect, $r = .78$, 95% CI [.77, .80]. The correlation between gender-career OB and age OB was a large effect, $r = .66$, 95% CI [.58, .89]. The correlation between gender-career G and age G was not different from 0, $r = .08$, 95% CI [−.03, .16]. These results are summarized in the bottom portion of Table 2.

Summary and Discussion

The purpose of Study 2 was to replicate the pattern of domain-generality that emerged for the D parameter and clarify the pattern of results for the AC, OB, and G parameters. Mirroring the results of Study 1, the magnitude of D correlations in Study 2 was large, ranging from .67 to .78. As such, D appears to be a robust, relatively non-attitudinal process. The nature of AC was clarified greatly in Study 2, with average correlations ranging from large (.50) when conceptual overlap was high to small (.20) when conceptual overlap was low. As such, AC appears to be a relatively attitude-specific process. The nature of OB was also clarified greatly in Study 2. With large correlations, ranging from .67 to .85, OB appears to be a relatively non-attitudinal process. The least clear results emerged for G. With a moderate correlation (.21) when conceptual overlap was high, but no correlation when overlap was moderate, G appears to be an attitude-specific process (the G parameters represent different processes on the low overlap IATs and, therefore, cannot be meaningfully compared). It is important to note that, across all six studies, the majority of the correlation coefficients for all Quad processes were greater than 0 but less than 1. In other words, a parameter derived from one IAT neither accounts for all nor none of the variances in the same parameter on a different IAT. This suggests that each of the Quad processes possesses varying degrees of attitude-specificity and domain-generality.

General Discussion

The IAT was designed to measure associations specific to a given attitude or content domain, and has accumulated an impressive record of validity across a growing number of fields (e.g., Greenwald et al., 2009). Previous work has shown that IAT performance is influenced by both associative and non-associative processes (e.g., Sherman et al., 2008), but until now, the extent to which these processes were attitude-specific versus content-general (non-attitudinal) had remained largely unexamined. The present research addresses this theoretical gap by demonstrating that IAT performance reflects the influence of both relatively attitude-specific processes and relatively general, non-attitudinal processes.

Detection (D) is a relatively domain-general, non-attitudinal process. Across all studies, the magnitude of D parameter correlations was large, regardless of conceptual overlap. D reflects a relatively controlled accuracy-oriented process that discriminates between correct and incorrect responses; participants who can determine the correct response in one domain can do so in others. This may reflect individual differences in the motivation and/or ability to focus on whatever task these individuals are doing.

AC is a relatively attitude-specific process that reflects associations activated by the content of a given IAT. The magnitude of AC correlations was larger across IATs with high conceptual overlap than with less conceptual overlap. However, even in the low conceptual overlap conditions, AC parameter correlations were greater than 0, suggesting a degree of domain-generality as well. This may reflect individual differences, such as a general Need for Evaluation (Jarvis & Petty, 1996) that may influence the strength of associations across domains. It may also indicate some as yet undetermined source of task variance that is picked up by the AC parameter. Future research should further investigate these possibilities.

OB is a relatively domain-general, non-attitudinal process. Indeed, the large magnitude of OB correlations across Study 2 and most of Study 1 was roughly equal to that of the D correlations. The variability in OB correlations across Study 1 is perhaps due, in part, to the structure of the Quad model, which estimates OB only from responses to target stimuli on incompatible trials (i.e., on 25% of the trials). As such, large samples are necessary to identify its effects. OB is an inhibitory process that constrains the expression of an activated association when it conflicts with a contextually accurate response (Sherman et al., 2008). Because inhibition is often conceptualized as a domain-general ability, the present research situates OB within a larger constellation of executive function (Friedman & Miyake, 2004).

G appears to be a relatively attitude-specific process. The small-to-medium magnitude of correlations across high overlap tests but near-zero correlations across moderate overlap tests suggests that G is sensitive to conceptual overlap. (It is not possible to interpret the correlations across low overlap tests, as these G parameters represent qualitatively different processes). G reflects the influence of general guessing or response biases that may influence behavior in the absence of other available guides to responses, and appears to be an attitude-specific non-associative process.

Implications

Given the ubiquity of the IAT, this research has important implications for interpretation of IAT scores, individual differences in IAT scores, malleability of IAT scores (and the development of interventions), and for demonstrations of IAT scores predicting behavior. Because some of the processes that contribute to IAT performance are not related to the specific construct being measured, individual differences in IAT scores and the malleability of those scores may reflect differences in processes that are not content-specific. This, in turn, suggests more nuanced interpretations of previous research. Take, for example, the previously discussed findings of Gonsalkorale, Sherman, and Klauer (2009) that older people's greater racial IAT bias is related to the decreased ability to regulate biased associations (OB). The present research supports an interpretation of these results as reflecting age-related deficits in general cognitive function that may have little to do with racial bias, per se.

Not only does the present research suggest new interpretations of previous research, but it can also guide the development of bias-reduction interventions. Depending on the strengths or deficits of the target population, or any of a variety of constraints or goals, future bias-reduction interventions can be tailored to focus on attitude-specific processes, non-attitudinal processes, or any combination of processes. Moreover, it is possible that interventions that target the non-attitudinal processes that influence IAT performance could have important behavioral consequences external to the IAT. For example, Calanchini, Gonsalkorale, Sherman, and Klauer (2013) demonstrated that D is responsive to training: Participants who completed a counter-prejudicial training task subsequently demonstrated less IAT bias and had higher activation of D than control participants. Whether or not this training carries over to influence accuracy detection processes beyond the IAT to other tasks and behaviors remains an open question to be explored. Likewise, it remains an open question whether domain-general training that influences processes such as D or OB will subsequently influence performance in a specific domain. Previous research (Gonsalkorale, Sherman, Allen, Klauer, & Amodio, 2011) has shown that D is related to correct identification of weapons and tools in the Weapons Identification Task (Payne, 2001). In combination with the present research, this suggests that training police officers, for example, to improve their general object identification skills may influence their ability to accurately identify the presence of a weapon in the field. Further research should explore this possibility. Another open question remains whether processes such as D and OB, which are both non-associative and non-attitudinal but are otherwise qualitatively distinct, are equally responsive to interventions or respond to different interventions differentially. Of course, the same point applies to attitude-specific processes such as AC and G, and future research may reveal important differences in malleability within attitude-specific and non-attitudinal processes.

The present research also highlights the utility of applying multinomial models to better isolate the attitudinal components of IAT performance. Indeed, one interpretation of the present research might be that the AC parameter provides a cleaner measure of attitude than the IAT D score. However, we would hesitate to make such a strong claim without the support of careful validation studies. Regardless of which measure is a relatively more pure measure of attitude, the IAT D score may be a better predictor of behaviors that also involve domain-general processes such as D and OB. The present research by no means impugns the predictive validity of the IAT, but rather provides a more nuanced understanding of some of the processes that contribute to IAT performance. To the extent that the same non-attitudinal processes that influence IAT performance also affect behaviors of interest, then it may be possible to observe IAT performance predicting attitude-relevant behavior without the involvement of the attitude, per se. Future research should explore the extent to which these processes, both attitude-specific and non-attitudinal, contribute to behavior.

Conclusion

The Quad model does not provide an exhaustive account of IAT performance, which is certainly influenced by processes that are not included in the model, including recoding (e.g., Chang & Mitchell, 2011; De Houwer et al., 2005; Kinoshita & Peek-O'Leary, 2005, 2006; Meissner & Rothermund, 2013; Rothermund et al., 2009; Rothermund & Wentura, 2001, 2004; Rothermund et al., 2005), task-set shifts and task-set simplification (e.g., Mierke & Klauer, 2001, 2003; Klauer et al., 2010), and speed-accuracy tradeoffs (e.g., Brendl et al., 2001; Klauer et al., 2007). Klauer and colleagues have already demonstrated method variance in IAT performance related to task switching ability (Klauer et al., 2010; Mierke & Klauer, 2001, 2003). Future research examining the domain specificity of the other component processes would also represent important contributions to our understanding of the IAT and attitudinal responses, more generally. Finally, future research may also apply the type of analysis presented in this article to other implicit measures of attitude, such as evaluative priming (Fazio, Jackson, Dunton, & Williams, 1995) or the Affect Misattribution Procedure (AMP; Payne, Cheng, Govorun, & Stewart, 2005). We suspect that all implicit measures reflect some degree of influence from non-attitudinal processes that may be consequential to understanding effects obtained with those measures.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

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

Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, *83*, 380-393.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, *6*, 57-86.

Beer, J. S., Stallen, M., Lombardo, M. V., Gonsalkorale, K., Cunningham, W. A., & Sherman, J. W. (2008). The Quadruple Process model approach to examining the neural underpinnings of prejudice. *NeuroImage*, *43*, 775-783.

Brendl, C. M., Markman, A. B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology*, *81*, 760-773.

Calanchini, J., Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2013). Counter-prejudicial training reduces activation of biased associations and enhances response monitoring. *European Journal of Social Psychology*, *43*, 321-325.

Chang, B. P. I., & Mitchell, C. (2011). Discriminating between the effects of valence and salience in the implicit association test. *The Quarterly Journal of Experimental Psychology*, *64*, 2251-2275.

Cohen, J. (Ed.). (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Connelly, S. L., Hasher, L., & Zacks, R. T. (1991). Age and reading: The impact of distraction. *Psychology and Aging*, *6*, 533-541.

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The Quad Model of implicit task performance. *Journal of Personality and Social Psychology*, *89*, 469-487.

De Houwer, J., Geldof, T., & De Bruycker, E. (2005). The Implicit Association Test as a general measure of similarity. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *59*, 228-239.

Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin*, *21*, 1139-1150.

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

Forehand, M. R., & Perkins, A. (2005). Implicit assimilation and explicit contrast: A set/reset model of response to celebrity voice-overs. *Journal of Consumer Research*, *32*, 435-441.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, *133*, 101-135.

Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers [Special issue: Developments in psychological measurement and assessment]. *Canadian Psychology/Psychologie canadienne*, *50*, 141-150.

Gawronski, B., Ehrenberg, K., Banse, R., Zukova, J., & Klauer, K. C. (2003). It's in the mind of the beholder: The impact of stereotypic associations on category-based and individuating impression formation. *Journal of Experimental Social Psychology*, *39*, 16-30.

Gawronski, B., Peters, K. R., Brochu, P. M., & Strack, F. (2008). Understanding the relations between different forms of racial prejudice: A cognitive consistency perspective. *Personality and Social Psychology Bulletin*, *34*, 648-665.

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

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology*, *45*, 410-414.

Gonsalkorale, K., von Hippel, W., Sherman, J. W., & Klauer, K. C. (2009). Bias and regulation of bias in intergroup interactions: Implicit attitudes toward Muslims and interaction quality. *Journal of Experimental Social Psychology*, *45*, 161-166.

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, *79*, 1022-1038.

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.

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.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193-225). San Diego, CA: Academic Press.

Jarvis, W. B. G., & Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology*, *70*, 172-194.

Kinoshita, S., & Peek-O'Leary, M. (2005). Does the compatibility effect in the race Implicit Association Test reflect familiarity or affect? *Psychonomic Bulletin & Review*, *12*, 442-452.

Kinoshita, S., & Peek-O'Leary, M. (2006). Two bases of the compatibility effect in the Implicit Association Test (IAT). *The Quarterly Journal of Experimental Psychology*, *59*, 2102-2120.

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika*, *75*, 70-98.

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

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

Maddox, K. B. (2004). Perspectives on racial phenotypicality bias. *Personality and Social Psychology Review*, *8*, 383-401.

Meissner, F., & Rothermund, K. (2013). Estimating the contributions of associations and recoding in the Implicit Association Test: The ReAL model for the IAT. *Journal of Personality and Social Psychology*, *104*, 45-69.

Mierke, J., & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Zeitschrift für Experimentelle Psychologie*, *48*, 107-122.

Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology*, *85*, 1180-1192.

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

Nosek, B. A., Banaji, M., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, *6*, 101-115.

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*, 181-192.

Payne, B. K., Cheng, S. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*, 277-293.

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). *The Quarterly Journal of Experimental Psychology*, *62*, 84-98.

Rothermund, K., & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test (IAT). *Zeitschrift für Experimentelle Psychologie*, *48*, 94-106.

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

Rothermund, K., Wentura, D., & De Houwer, J. (2005). Validity of the salience asymmetry account of the Implicit Association Test: Reply to Greenwald, Nosek, Banaji, and Klauer (2005). *Journal of Experimental Psychology: General*, *134*, 426-430.

Sherman, J. W. (2006). On building a better process model: It's not only how many, but which ones and by which means? *Psychological Inquiry*, *17*, 173-184.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review*, *115*, 314-335.

Teachman, B. A., Gregg, A. P., & Woody, S. R. (2001). Implicit associations for fear-relevant stimuli among individuals with snake and spider fears. *Journal of Abnormal Psychology*, *110*, 226-235.

Wiers, R. W., Houben, K., & de Kraker, J. (2007). Implicit cocaine associations in active cocaine users and controls. *Addictive Behaviors*, *32*, 1284-1289.

Wiers, R. W., Van Woerden, N., Smulders, F. T. Y., & De Jong, P. J. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology*, *111*, 648-658.