Can the Implicit Association Test Serve as a Valid Measure of Automatic Cognition? A Response to Schimmack (2021)

Benedek Kurdi[1] , **Kate A. Ratliff[2], and William A. Cunningham[3]**

[1]Department of Psychology, Cornell University; [2]Department of Psychology, University of Florida; and [3]Department of Psychology, University of Toronto

Abstract

Much of human thought, feeling, and behavior unfolds automatically. Indirect measures of cognition capture such processes by observing responding under corresponding conditions (e.g., lack of intention or control). The Implicit Association Test (IAT) is one such measure. The IAT indexes the strength of association between categories such as "planes" and "trains" and attributes such as "fast" and "slow" by comparing response latencies across two sorting tasks (planes–fast/trains–slow vs. trains–fast/planes–slow). Relying on a reanalysis of multitrait–multimethod (MTMM) studies, Schimmack (this issue, p. 396) argues that the IAT and direct measures of cognition, for example, Likert scales, can serve as indicators of the same latent construct, thereby purportedly undermining the validity of the IAT as a measure of individual differences in automatic cognition. Here we note the compatibility of Schimmack's empirical findings with a range of existing theoretical perspectives and the importance of considering evidence beyond MTMM approaches to establishing construct validity. Depending on the nature of the study, different standards of validity may apply to each use of the IAT; however, the evidence presented by Schimmack is easily reconcilable with the potential of the IAT to serve as a valid measure of automatic processes in human cognition, including in individual-difference contexts.

Humans have a remarkable ability for reflective thought, which enables us to write poetry, prove theorems in geometry, and discuss with each other the nature of the mind. At the same time, considerable consensus in the cognitive sciences exists that, in addition to controlled processes of cognition involved in reflective thought, information can also be activated relatively automatically in the human mind. After decades of research on automatic processes using myriad other measures (e.g., Carlston & Skowronski, 1994; Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Meyer & Schvaneveldt, 1971; Neely, 1976; Nuttin, 1985; Stroop, 1935), the Implicit Association Test (IAT) was introduced by Greenwald, McGhee, and Schwartz (1998) as a novel measure of relatively automatic information activation.

Like many other indirect measures,[1] the IAT consists of a series of sorting trials. These sorting trials are completed in two critical blocks. For instance, in an IAT designed to provide a measure of the relative association[2] of the categories "planes" and "trains" with the attributes "fast" and "slow," participants first use the same response key to sort stimuli representing the category "planes" (e.g., images of the Boeing 737-800, the Airbus A320, the Bombardier CRJ700, and the Embraer E175) and the attribute "fast" (e.g., synonyms such as *fast, quick, rapid*, and *speedy*) and a different response key to sort stimuli representing the category "trains" (e.g., images of the GE P40DC, the EMD GP38H-3, the Siemens Sprinter ACS-64, and the MPI GP15D) and the attribute "slow" (e.g., synonyms such as *plodding, slow-moving, sluggish*, and *tardy*). In a second critical block the assignment of categories to attributes is reversed, with "planes" and "slow" mapped onto the same response key and "trains" and "fast" mapped onto a different response key. The trials

Corresponding Author:
Benedek Kurdi, Department of Psychology, Cornell University
E-mail: bk493@cornell.edu

themselves are typically preceded by the instruction that participants should go as quickly as possible without making too many mistakes, thus creating a certain degree of time pressure. Crucially, participants are not asked to use or reflect on their beliefs about planes, trains, or any other mental content while completing the task; the effect emerges in the absence of such deliberate processes.

The idea underlying most scoring algorithms used to evaluate responding on the IAT, including the most frequently used improved scoring algorithm (Greenwald, Nosek, & Banaji, 2003), is fairly simple: If planes and trains do not differ from each other in terms of the extent to which they are automatically associated with notions of speed over slowness, then, all other things being equal, mean response latencies across the two critical blocks (i.e., the planes–fast/trains–slow block and the trains–fast/planes–slow block) should be the same. And if the speed and accuracy of responding across the two critical blocks differs from each other, such differences should be indicative of the fact that either planes or trains are automatically evaluated to be faster. Moreover, the extent of the difference in response latencies can be seen as an index of the strength of the difference in relatively automatic responding: Assuming identical standard deviations, a 400-ms difference in speed across the two critical blocks should be viewed as suggesting a stronger difference in association than a 40-ms difference.

Schimmack's Critique of the IAT

Early critiques of the IAT (e.g., Arkes & Tetlock, 2004; Blanton et al., 2009; Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013) were predicated on the idea that response-time differences emerging from simple button presses cannot possibly reveal the operation of cognitive processes that give rise to the kinds of high-level phenomena of interest to social psychology, including attitudes, beliefs, and self-concept. In this issue, Schimmack (2021) relies on a reanalysis of five existing studies (Axt, 2018; Bar-Anan & Vianello, 2018; Cunningham, Preacher, & Banaji, 2001; Falk, Heine, Takemura, Zhang, & Hsu, 2014; Greenwald, Smith, Sriram, Bar-Anan, & Nosek, 2009) to provide a fundamentally different kind of criticism of the validity of the IAT.

To offer a brief summary, the reanalyses reported by Schimmack find high correlations between relatively indirect (automatic) measures of mental content, as indexed by the IAT, and relatively direct (controlled) measures of mental content, as indexed by a variety of self-report scales. In fact, the statistical evidence presented in Schimmack's article is compatible with the possibility that the IAT and self-report measures may reflect the same underlying latent construct (i.e., attitude or evaluation) rather than the former reflecting implicit attitudes and the latter reflecting explicit attitudes, each stored separately in memory. Far from being a collection of random numbers, Schimmack's analyses provide evidence that the IAT can be highly associated with evaluative constructs as assessed by direct measures; thus, the crux of his argument, diametrically opposed to the early critiques cited above, is that direct and indirect measures of cognition may be too highly associated with each other.

Here we do not wish to comment on whether the analyses reported by Schimmack represent the optimal way to approach these data or, alternatively, whether the models chosen by the original authors whose data he reanalyzed may have been more appropriate. Competing statistical models may reveal different aspects of the structure of the same data; any model merely suggests the plausibility of potential relationships. Therefore, for the purposes of this commentary, we accept the analyses presented by Schimmack at face value and ask what it means for the validity and usefulness of the IAT as well as contemporary theories of attitudes and social cognition if indirect and direct measures can, indeed, reflect, at least in part, the same underlying latent construct.[3]

We begin our commentary with a strong point of agreement: Like Schimmack, we believe that discussions about the validity of the IAT, like discussions about the validity of the Likert scale, are not particularly meaningful. Rather, the emphasis should be on evaluating the validity of each specific instantiation of the test. Notably, standards of validity will differ considerably depending on the intended use of the specific IAT. We then devote some space to the concepts of attitude and automaticity, as commonly used and understood in contemporary literature. We do so because, although Schimmack's article does not cast any doubt on the ability of IATs to index attitudes or to do so in an automatic fashion, we believe that a meaningful debate about Schimmack's more specific argument presupposes a precise understanding of these central constructs.

Most importantly, the bulk of our commentary is devoted to a discussion of theoretical advances in implicit social-cognition work over the past 20 years. Specifically, we view Schimmack's main finding, according to which direct and indirect measures of attitude can be closely, and sometimes even very closely, associated with each other, as not at all incompatible with the way that many social-cognition researchers have thought about the construct of (implicit) evaluation.[4] Specifically, numerous existing theories have advanced the view that indirect and direct measures of cognition may index the same, or at least largely overlapping, representations (Cunningham & Zelazo, 2007; De Houwer, 2014; De Houwer, Van Dessel, & Moran, 2020; Fazio, 2007; Kruglanski & Gigerenzer, 2011; Van Bavel, Xiao,

& Cunningham, 2012). At the same time, we highlight the need for further integration between individual-difference and experimental approaches to implicit social-cognition work, in line with Schimmack's reasoning and reiterating previous warnings to the same effect (e.g., Kurdi & Banaji, 2017; Payne, Vuletich, & Lundberg, 2017b).

Finally, to situate Schimmack's argument in the broader context of IAT research, we also point out that the question of whether the IAT provides a valid indication of individual differences in implicit cognition, which constitutes the sole focus of the original article, is irrelevant to many uses of the test. Such applications include, among others, experimental studies probing the formation of novel attitudes and stereotypes as well as changes in existing ones, studies seeking to predict behavior with the highest possible degree of precision, and studies using the IAT situated at the regional, rather than individual, level of analysis.

Are Discussions About the Validity of "the IAT" Meaningful?

To start with a strong point of convergence, we could not agree more with Schimmack's position that "the IAT is a method, just like Likert scales are a method, and it is impossible to say that a method is valid" (p. 398). We view arguments about whether the IAT is generally valid as tantamount to discussions about the validity of blood-pressure meters, kitchen scales, and tape measures. Perhaps the idea of "Implicit Association Test" exists somewhere in a Platonic realm of forms along with the ideas of "blood-pressure meter," "kitchen scale," and "tape measure"; however, for the sake of empirical research, only specific IATs used by specific researchers for specific purposes are of interest. It is impossible to make claims about the validity of an entire family of measures, including the IAT, at this level of generality.

Crucially, just as many specific blood-pressure meters exist and the $10 knockoff might not work as well as the latest $100 model, IATs also come in countless different flavors and versions: In addition to the White/Black–good/bad attitude IAT, which has received more attention than perhaps any other instantiation of the test, IATs have been used to investigate associations in a vast number of domains, including between the categories "Dublin life" and "country life" and the attributes "positive" and "negative" (Barnes-Holmes, Waldron, Barnes-Holmes, & Stewart, 2009), the categories "alcohol" and "water" and the attributes "approach" and "avoid" (Lindgren, Westgate, Kilmer, Kaysen, & Teachman, 2012), and the categories "me" and "not me" and the attributes "anxious" and "calm" (van Harmelen et al., 2010).

The plane/train–fast/slow IAT hypothetically described above may not yet exist, but it could be implemented and administered within a matter of hours by anyone with access to an Internet-enabled computer and some modest coding skills. Whether it would be reasonable to do so is, of course, a different issue. On a related note, whether the IAT can be meaningfully stated to be "in search of *a* [emphasis added] construct" (as it is in the title of Schimmack's article) seems questionable to us. As the examples cited above demonstrate, the IAT can be used to index a variety of different constructs, including attitudes, stereotypes, beliefs, self-esteem, self-concept, and others across countless different domains—and to do so without asking participants to make an intentional judgment using a Likert scale, feeling thermometer, or similar measure. Indeed, Schimmack's own analyses suggest that those IATs that he investigated do, in fact, index meaningful latent constructs, although those latent constructs are not dissociated from the latent constructs indexed by parallel direct measures.

Crucially, as a function of the specific targets investigated, the relationship between explicit and implicit constructs and direct and indirect measures may differ both across and within the bounds of the constructs of attitude, stereotype, belief, self-esteem, and self-concept. For instance, direct and indirect measures pertaining to the same construct (e.g., evaluation) may be more or less highly associated with each other depending on the object of the evaluation (e.g., political preferences vs. gender attitudes; Nosek, 2005). Moreover, direct and indirect measures of one construct (e.g., self-esteem) may generally be more dissociated from each other, and thus more likely to load on distinct latent variables, than direct and indirect measures of a different construct (e.g., beliefs; see Irving & Smith, 2020).

Can IATs Validly Index Attitudes?

IATs are perhaps most frequently used to measure attitudes (i.e., evaluations of stimuli along a positive–negative continuum; Eagly & Chaiken, 1993). As discussed in more detail below, validating a measure requires a precise theoretical understanding, or at least a well-defined theoretical idea, of the nature of the latent construct that the measure is proposed to index. In this context, it should be noted that Schimmack's claim according to which "most researchers regard the IAT as a valid measure of enduring attitudes that vary across individuals" (p. 397) does not reflect the theoretical consensus in the field. Specifically, it is accurate that attitude representations are not usually seen as completely ephemeral (but see Schwarz, 2007). Nonetheless, most attitude researchers do not subscribe to the

view that attitudes are enduring properties of individuals akin to their blood type or eye color, as suggested by Schimmack. If this were the case, the entire, highly voluminous, literature on attitude change would be an exercise in futility.

Rather, the overwhelming theoretical consensus in the community of attitude researchers, dating back to the view expressed by Mischel (1968), is that attitudes emerge from an interaction of persons and situations. In fact, this view is put forth in several articles cited by Schimmack himself (incorrectly) as supporting the proposition that attitudes are stable over time but vary across individuals (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009; Gawronski & Bodenhausen, 2017; Kurdi & Banaji, 2017; Rae & Greenwald, 2017). As expressed perhaps most succinctly by Rae and Greenwald (2017), "very few psychologists view human social behavior as being governed exclusively by either person or situation variations. Most consider it reasonable to ask how person and situation variations jointly influence social behavior" (p. 299). Indeed, far from arguing that "there are no stable attributes that influence performance on the IAT" (Schimmack, 2021, p. 398), Payne et al. (2017b) merely wish to shift the emphasis away from what they see as an excessive focus on personal factors and toward a fuller consideration of the social contexts in which processes of evaluation unfold.

As evidenced by the lively debate that followed the publication of the theoretical article by Payne, Vuletich, and Lundberg (2017a), the jury is still out on whether variation in responding on the IAT mostly reflects individual differences or mostly reflects the effects of the situation and, ultimately, whether the attitude construct itself should be seen as primarily affected by variation at the level of individuals or at the level of contexts. Alternatively, this juxtaposition may not be meaningful at all if the effects of both sources of variation are multiplicative rather than additive. It seems clear that without studies that measure the same individuals across multiple contexts, this question will remain impossible to answer.

Crucially for the current purposes, whether attitudes reflect mainly the person, mainly the situation, or some combination of both does not, in and of itself, validate or invalidate the IAT as a measure of attitudes.[5] Rather, evidence on the validity of an IAT can be interpreted only within a coherent theoretical framework, whatever that theoretical framework might be.[6] We recognize that Schimmack's article does not cast any doubt on the potential of the IAT to serve as a valid measure of attitudes. If anything, the findings reported in the article may be seen as supporting the validity of the IAT as a measure of attitudes given that it loads on the same latent factor as direct measures, even in the absence of

any shared method variance. At the same time, a meaningful answer to the central question of the article—whether IATs can be valid measures of individual differences in implicit attitudes that are dissociable from their explicit counterparts—presupposes a precise theoretical understanding of the attitude construct.

Can IATs Validly Index Automatic Cognition?

Schimmack's evidence exclusively concerns uses of the IAT as a measure of individual differences. Nonetheless, his conclusion that the IAT "is not a window into people's unconscious feelings, cognitions, or attitudes" (Schimmack, 2021, p. 412) may easily be misinterpreted as a far more general claim that IATs cannot validly index automatic processes in human cognition.[7] But Schimmack's article could not form the basis for such a sweeping claim given that it investigates only a minuscule subset of the IAT literature, none of which is explicitly concerned with establishing the automaticity conditions of the measure. That said, Schimmack's more specific claim about the use of the IAT as an individual-difference measure also cannot be appropriately evaluated without a precise theoretical understanding of the nature of automaticity.

In this context, it is important to point out that even though "implicit bias" and "unconscious bias" are often used synonymously in popular discourse, in the relevant literature the terms "implicit" and "unconscious" are not commonly seen as interchangeable. Nearly 3 decades ago Bargh (1994) made the influential observation that automaticity is not a unitary construct (see also De Houwer et al., 2009; Gawronski, 2019). Rather, in addition to conscious awareness (which itself has multiple facets), the automatic nature of a cognitive process is also determined by additional features such as intention, efficiency, and control. In other words, a conscious process can be automatic to the degree that it (a) unfolds in the absence of the person's volitional decision to initiate it (intention), (b) is not affected by concurrent processes that depend on working memory (efficiency), or (c) cannot be stopped voluntarily (control).

Indeed, evidence suggests that participants may, to some extent, be aware of the latent construct indexed by some IATs (Hahn & Gawronski, 2019; Hahn, Judd, Hirsh, & Blair, 2014). However, to our knowledge, no implicit social-cognition researcher has understood such evidence to cast any doubt on the automatic nature of the IAT because the IAT and direct measures clearly index their respective underlying latent constructs under different automaticity conditions (De Houwer et al., 2009): On Likert scales and feeling thermometers, participants provide intentional judgments of attitude, stereotype, or self-concept. By contrast, on the IAT, participants complete

a combined sorting task, with attitudes, stereotypes, or self-concept indirectly inferred from patterns of response latencies and errors in the absence of any intention on the participant's part to provide judgments on such mental content. Moreover, unlike direct measures, IATs create suboptimal conditions in that participants are typically instructed to respond as fast as they can.

Together, these features of the IAT make it a measure of automatic processes in human cognition even if evidence of awareness can be provided for some specific tests under some specific conditions (e.g., Hahn et al., 2014; Hahn & Gawronski, 2019) and, importantly, even if specific IATs are sometimes highly correlated with parallel direct measures (e.g., Schimmack, 2021). Furthermore, the fact that participants are above chance in predicting the rank ordering of their IAT scores on a certain number of specific tests does not imply that all participants are aware of all of their (implicit) attitudes all of the time. And, needless to say, unconscious attitudes, stereotypes, and self-concept may very well exist even if they are not measured by the IAT. Finally, even a certain degree of awareness of the existence or the strength of one's IAT scores does not necessarily imply that implicit attitudes are fully available to conscious introspection. For instance, one might be aware of one's implicit attitudes, and yet implicit attitudes may affect judgments and behaviors in a manner that bypasses conscious awareness.

Can IATs Validly Index Individual Differences in Implicit Attitudes?

Notably, the evidence presented by Schimmack cannot speak to the general potential of the IAT to validly index attitudes or automatic processes in human cognition. Rather, throughout most of his article, Schimmack seems to advance the considerably more specific claim that, in the context of work on individual differences (which constitutes only a subset of IAT research), the IAT and direct measures of cognition can index the same underlying latent construct. That is, rather than the IAT indexing implicit evaluations and Likert scales indexing explicit evaluations, both measures can index the same latent construct of evaluation. As mentioned above, in making this argument, Schimmack relies on reanalyses of existing studies using a multitrait–multimethod (MTMM) framework (Axt, 2018; Bar-Anan & Vianello, 2018; Cunningham et al., 2001; Falk et al., 2014; Greenwald et al., 2009). Unlike Schimmack, we do not believe that these findings call into doubt the potential of the IAT to serve as a valid measure of individual differences in automatic cognition. They seem difficult to reconcile with a theoretical view positing that explicit and implicit evaluations emerge from qualitatively different (noninteracting) processes, systems, or

representations; however, the validity of the IAT does not presuppose the accuracy of this theoretical view.

Crucially, the use of the MTMM matrix to investigate patterns of association and dissociation among different measures representing the same and different constructs, as done by Schimmack and by countless other investigators before him, is meaningful only within the context of a well-formulated theory about how those constructs are expected to relate to each other (Campbell & Fiske, 1959). If a researcher's theory posits that different labels being used to refer to self-esteem and narcissism are simply a happenstance of natural language, they will not be particularly surprised to find high correlations or even complete redundancy between measures of self-esteem and measures of narcissism (or, using more contemporary analytic techniques, the latent traits indicated by them in a structural equation modeling framework). The same observation also applies to implicit social cognition: High correlations between parallel direct and indirect measures, for instance, a Likert scale and an IAT measuring relative judgments of how fast planes are versus trains, should be seen as surprising only to the degree that a substantive theory of the domain predicts that responding on direct and indirect measures should index different underlying latent constructs, in this particular case explicit and implicit beliefs about the speed of different means of transportation.

Although relatively strict dual-process theories positing that direct and indirect measures assess fundamentally different underlying representations certainly exist (McConnell & Rydell, 2014; Rydell & McConnell, 2006; Smith & DeCoster, 2000; Strack & Deutsch, 2004), there is by no means consensus in the field that these theories are accurate (Cunningham & Zelazo, 2007; De Houwer, 2014; De Houwer et al., 2020; Fazio, 2007; Kruglanski & Gigerenzer, 2011; Van Bavel et al., 2012). As explained below, from the perspective of theories that assume that the same, or at least largely overlapping, processes underlie responding on direct and indirect measures of evaluation and belief, the results reported by Schimmack are not particularly surprising. Whether dual-process theories of attitudes and higher-order cognition in general are accurate is certainly an important issue with which to grapple; however, evidence against such theories is not evidence against the validity of the IAT.

Alternatives to dual-process theories of implicit cognition

Although it may appear that there is consensus regarding the cognitive processes undergirding automatic evaluation and cognition, there remains significant debate within social psychology regarding underlying mechanisms.

Indeed, much of social-cognition research over the past 2 decades has tried to understand the nature of automatic cognition and to find the most plausible interpretation of changes in indirect measures. Although the dual-attitude structure proposed in the review by Greenwald and Banaji (1995), which served as the conceptual framework for the new method introduced by Greenwald et al. (1998), has been an influential model in the field (see also McConnell & Rydell, 2014; Rydell & McConnell, 2006; Smith & DeCoster, 2000; Strack & Deutsch, 2004), it is by no means the only or even the clearly dominant theoretical perspective today. In fact, the originators of the IAT recently argued that the distinction between explicit and implicit cognition should a priori be seen as situated at the level of measures rather than mental constructs, thus leaving the issue of association versus dissociation to empirical research (Greenwald & Banaji, 2017).

In an influential early alternative to the dual-process perspective, Fazio (2007) argues that there is a single attitude construct in memory, represented as a link between an attitude object and an evaluation (e.g., plane–bad, train–good).[8] In this view, both direct measures, such as Likert scales, and indirect measures, such as IATs, index the same underlying latent construct of attitude. Thus, given that the theory does not posit the existence of a separate implicit attitude, high levels of association between responding on direct and indirect measures are by no means unexpected. At the other end of the spectrum, a more recent set of theories by De Houwer and colleagues (De Houwer, 2014; De Houwer et al., 2020; Hughes, Barnes-Holmes, & De Houwer, 2011) suggest that, rather than associative representations, propositional information underlies responding on both direct and indirect measures. For instance, responding on the plane/train attitude IAT may be driven by the proposition that planes are bad and trains are good. Finally, the iterative reprocessing framework proposed by Cunningham and colleagues (Cunningham & Zelazo, 2007; Cunningham, Zelazo, Packer, & Van Bavel, 2007; Van Bavel et al., 2012) assumes that stimulus evaluations emerge from a series of processing steps in the course of which evaluations are gradually adjusted in light of contextual and motivational information, with no strict separation between the constructs reflected by direct and indirect measures.

At the same time, each of these theories can also account for dissociations between direct and indirect measures without hypothesizing the existence of distinct implicit and explicit constructs in memory. Fazio (2007) notes that responding on direct measures can be influenced by a host of nonattitudinal factors, including the motivation to appear nonprejudiced, which may result in an attenuation of the correlation between direct and indirect measures of attitude. According to De Houwer (2014), given that direct and indirect measures differ from each other in terms of automaticity features (De Houwer et al., 2009), a different set of propositions may be activated when responding on direct as opposed to indirect measures of cognition. Specifically, the use of indirect measures may result in "quick and dirty" reasoning that does not fully conform to the rules of propositional logic. Finally, under the framework proposed by Cunningham et al. (2007), responding on direct and indirect measures may differ from each other because the automaticity conditions imposed by indirect measures cut short the reprocessing steps that would unfold to a fuller extent under the more optimal conditions afforded by direct measures.

Any attempt to arbitrate between these specific theoretical proposals and their dual-process alternatives is well beyond the scope of this commentary. Nonetheless, this quick overview should make it clear that, in sharp contrast with the theoretical landscape that characterized social-cognition research 20 years ago, dual-process theories are by no means the only dominant—or perhaps not even the clearly dominant—available theoretical perspective anymore. In addition, to the extent that theories posit the same, or at least partially overlapping, cognitive operations to drive responding on direct and indirect measures of evaluation and belief, the results of association reported by Schimmack are, if anything, more expected than results of dissociation.

Implications for individual differences

At the same time, it should be noted that the theories reviewed above have all been formulated within the framework of experimental social psychology and, as such, focus on the conditions under which new learning, as well as shifts to existing representations, should occur. Notably, we are not aware of any detailed discussion of the results that such theories would predict to emerge in MTMM investigations, which rely on patterns of correlation with other measures rather than theoretically predicted patterns of acquisition and change. In this context, we can note only that, sadly, the separation between experimental and correlational approaches is still alive and well after over 60 years of the initial warning about the need for greater integration across the "two disciplines of scientific psychology" (Cronbach, 1957). We see the article by Schimmack, as well as the recent theoretical contribution by Payne et al. (2017a), as reinforcing this warning.

However, despite the lack of specific existing discussions of the individual-difference context, it is possible to derive relevant predictions from all contemporary accounts of implicit cognition. Specifically, it seems clear that none of the theories reviewed above would unconditionally predict perfect association or perfect

dissociation in an individual-difference framework: Although they do not suggest that different memory representations underlie responding on direct and indirect attitude measures, each of these theories foresees the possibility that responding on direct measures may be modulated by processes not captured by indirect measures. Thus, under these views, patterns of association and dissociation can be investigated in a theoretically meaningful way only if such additional processes are appropriately taken into account (see also Gawronski, 2019; Kurdi et al., 2019).

For instance, as mentioned above, Fazio (2007) posits that responding on direct measures reflects both an underlying evaluative association with the attitude object and a host of other psychological processes, including nonevaluative knowledge about the attitude object, the motivation to appear nonprejudiced, and self-presentational concerns. Moreover, under some conditions, IATs may reveal knowledge that is not available to self-report because of a lack of prior elaboration. For instance, a participant may be surprised to learn that they associate the concept "why" with the concept "far" and the concept "how" with the concept "near" (in line with construal-level theory; Trope & Liberman, 2010). This result may be unexpected for the participant not because the associations revealed by this IAT are socially undesirable but rather because, unlike with social issues, everyday life does not usually afford many opportunities for elaborating on conceptual relationships of this kind.

In any case, according to accounts positing that direct measures reflect a mix of attitudinal and nonattitudinal processes, the data reanalyzed by Schimmack are insufficient for deciding whether direct and indirect measures reflect the same latent construct: From the perspective of such theories, any model that omits the nonattitudinal variables thought to contribute to responding on direct measures, including self-presentational concerns, prior elaboration, and others, should be seen as misspecified. Moreover, requiring unconditional dissociation to establish construct validity would lead to absurd consequences that preclude meaningful empirical investigation of the relationship between direct and indirect measures of attitude. For example, there is considerable contextual variation in the degree to which participants are motivated or able to monitor their automatic reactions to different attitude objects (Fazio, 1990). Under Schimmack's view, the IAT by definition could not serve as a valid measure of automatic cognition in those contexts in which more automatic and more deliberate reactions are aligned with each other.

Empirical evidence on dissociations

To summarize, the reanalyses reported by Schimmack seem to be relatively difficult to reconcile with a strict

separation of explicit and implicit evaluations, as postulated by traditional dual-process theories, and may be seen as evidence in favor of a view positing a considerable amount of overlap between the two. Crucially, as discussed above, we do not believe that the potential of the IAT to serve as a valid indirect measure of individual differences in attitudes hinges on the accuracy of the dual-process proposal. However, at the same time, for the sake of accuracy, it should be noted that plenty of evidence in favor of dissociations between direct and indirect measures exists. It is our view that any convincing theory of implicit social cognition must provide an account of why such dissociations arise, including, notably, in the context of attitudes toward social groups.

For instance, a recent meta-analysis by Kurdi et al. (2019) provided multiple demonstrations, using four different analytic approaches, that IATs are, on average, associated with measures of intergroup behavior above and beyond parallel direct measures of attitudes, stereotypes, and self-concept. Moreover, the effect sizes associated with the unique contributions of direct and indirect measures were found to be virtually identical. Similar patterns of incremental predictive validity have also been observed in other contexts, including psychopathology (e.g., Lindgren et al., 2016) and close relationships (e.g., Faure, Righetti, Seibel, & Hofmann, 2018).

Although these findings may not be theoretically conclusive, at the very least, they jointly indicate that, contrary to the limited evidence presented by Schimmack, direct and indirect measures can have meaningful nonoverlapping portions of variance. The conditions under which dissociations can occur, as well as the reasons for such dissociations, are being actively debated; nonetheless, at least tentatively, they can be seen as evidence in favor of a dual-process view (but see Van Dessel, Gawronski, & De Houwer, 2019). However, crucially, the potential of IATs to serve as valid measures of automatically revealed associations does not presuppose the accuracy of any specific theoretical proposal. Rather, different theoretical frameworks will result in different expectations about the patterns of association or dissociation that should emerge between IATs and direct measures of social cognition, often depending on the presence of some third variable.

Can IATs Serve as Valid Measures Beyond an Individual-Difference Context?

Schimmack (2021) opens with the observation that "relatively little is known about the construct validity of the IAT" (p. 396). This conclusion seems predicated on a view that equates the concept of construct validity with the concept of nomothetic span—that is, the relationship

of a measure with other measures hypothesized to reflect the same and different underlying latent constructs (Campbell & Fiske, 1959). This approach may be defensible to the degree that the focus is, as it is in Schimmack's article, exclusively on using an instrument to index individual differences. However, for the purposes of providing a more complete view of the field, it should be noted that much, perhaps even most, work relying on IATs uses versions of the test for other purposes, such as to measure group differences in experimental research, to improve the prediction of consequential outcomes, or to probe differences not across individuals but rather across geographic regions. Crucially, depending on the intended use of an instrument, different types of empirical data can be regarded to constitute convincing evidence of construct validity (Messick, 1995). Specifically, the MTMM matrix, which constitutes the sole focus of Schimmack's investigation, cannot provide evidence on validity beyond an individual-difference context.

Looking beyond nomothetic span

Since the inception of modern psychometrics, theoreticians and practitioners have recognized the importance of considering evidence on construct representation in addition to evidence on nomothetic span for establishing the validity of a measure (e.g., Cronbach & Meehl, 1955; Embretson, 1998; Messick, 1995; Sternberg, 1981; Whitely, 1983). Evidence on construct representation seeks to "identify . . . the theoretical mechanisms that underlie task performance" (Whitely, 1983, p. 180) and to "understand . . . the processes, strategies, and knowledge that persons use to solve items" (Embretson, 1998, p. 382). Under a relatively recent proposal by Borsboom, Mellenbergh, and van Heerden (2004), construct representation should, in fact, be seen as the only relevant source of evidence on construct validity. According to this view, the sole aim of construct validation should be to establish that (a) the phenomenon that the test is purported to measure exists and (b) it is causally responsible for changes in the outcome of the measurement procedure.

We do not wish to engage with the details of the debate on whether nomothetic span should be part of construct validity; nonetheless, several points made by Borsboom et al. (2004) seem, at the very least, fairly instructive. For instance, these authors point out that if every participant exhibits the same score on a measure, under a nomothetic-span approach, this measure is by definition invalid because it cannot be correlated with any other measure. However, from the perspective of construct representation, this claim is absurd: Suppose

that every single person were found to show a 400-ms average difference in response latency across the two critical blocks of our imaginary plane/train–fast/slow IAT. Such complete lack of variability does not seem to preclude the measure from validly indexing the level of automatically revealed association between the categories "plane" and "train" and the attributes "fast" and "slow."

At a minimum, it should be noted that nomothetic span and construct representation are independent of each other. Thus, when considering the construct validity of a measure, evidence on nomothetic span should be supplemented with evidence on construct representation. In other words, a measure can be demonstrated to be an excellent measure of individual differences even if the cognitive mechanisms that are causally responsible for responding on the measure are not well understood. Conversely, every psychological measure need not be a good measure of individual differences to be valid as a measure of an underlying latent construct. For instance, several classic indirect measures, such as the Stroop task (Stroop, 1935), have been shown to be ideally suited for experimental work precisely because they create consistently large effects without much variation across participants (Hedge, Powell, & Sumner, 2018). At the same time, such lack of meaningful variability at the individual level does not compromise the ability of a measure to index basic cognitive processes, such as response interference, at the group level.

Moreover, in some contexts, what construct is measured by an IAT and via what mechanisms may be almost entirely irrelevant. For instance, in contexts in which IATs are used to predict consequential outcomes that are themselves difficult to measure or whose occurrence should be avoided, such as suicide or nonsuicidal self-injury (e.g., Barnes et al., 2016; C. R. Glenn, Millner, Esposito, Porter, & Nock, 2019; J. J. Glenn et al., 2017), the only important aspect of the measure might be whether it is related to some criterion behavior above and beyond other known predictors. To the degree that the emphasis is on understanding how human cognition operates, establishing a bivariate or even multivariate relationship with another variable may not be particularly conclusive: Loevinger (1957) famously quipped that criterion validity "contributes no more to the science of psychology than rules for boiling an egg contribute to the science of chemistry" (p. 641). However, some eggs may be inherently important to boil; moreover, as discussed above, when issues of basic process are implicated, patterns of correlation, even within the context of the well-known and widely used MTMM matrix, are not usually seen as the sole source of relevant evidence on the validity of a measure.

Additional sources of evidence on validity of the IAT

A detailed review of the empirical evidence on the construct representation of IATs is beyond the scope of this commentary. However, we note that evidence on historical precursors of the IAT (e.g., Carlston & Skowronski, 1994; Fazio et al., 1986; Meyer & Schvaneveldt, 1971; Neely, 1976; Nuttin, 1985; Stroop, 1935), formal process models of task performance (e.g., Calanchini, Sherman, Klauer, & Lai, 2014; Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007; Meissner & Rothermund, 2013), theoretically relevant manipulations that modulate responding on IATs (e.g., Cone, Mann, & Ferguson, 2017; De Houwer et al., 2020; Gawronski & Bodenhausen, 2006, 2011), and known group differences (e.g., Barnes-Holmes et al., 2009; Lindgren et al., 2012; van Harmelen et al., 2010) should also be seen as part of the evidence on construct validity. Crucially, as pointed out above, even if every individual in the world showed the exact same difference in response latency across the two critical blocks, then, contrary to Schimmack's assertion, some IATs may still be able to provide a "window into the unconscious" (Schimmack, 2021, p. 412). Claims about operating conditions of a measure (e.g., lack of awareness or intentionality) should rest on careful experimental approaches and process modeling rather than studies examining a pattern of correlation with other measures.

Conclusion

Introduced by Greenwald et al. (1998) more than 2 decades ago, the IAT is still widely used as a measure of association between categories such as "train" and "plane" and attributes such as "slow" and "fast," as revealed under relatively automatic processing conditions. Early critiques of the measure (e.g., Arkes & Tetlock, 2004; Blanton et al., 2009; Oswald et al., 2013) were based on the objection that high-level cognition, such as the mental operations giving rise to attitudes, beliefs, and self-concept, requires controlled processing and conscious endorsement. As such, so the argument went, the presence of attitudes, beliefs, and self-concept cannot be inferred from a task that relies on the speed of button presses rather than on self-report.

Notably, Schimmack makes the opposite argument. His investigation found attitudes revealed by direct and indirect measures to be too similar to each other to be able to posit the existence of separate underlying representations. We see this as a remarkable development: From a task that some believed lacked the ability to tap constructs of central importance to psychological inquiry because it was too dissimilar from existing direct measures, the IAT has, over the past 20 years, morphed into a measure that is now seen by others to be too similar to direct measures to reveal anything interesting about the mind. Both of these observations cannot be accurate at the same time. And we are of the view that the truth is somewhere in the middle: The IAT is an index of attitudes, beliefs, and self-concept, as revealed under relatively automatic conditions. Sometimes such automatically revealed attitudes, beliefs, and self-concept are highly similar to their directly measured counterparts and sometimes they are not (Nosek, 2005, 2007). Current theories of social cognition recognize both possibilities.

It is our understanding that Schimmack's argument is not meant to apply to uses of the IAT other than as a measure of individual differences. However, whether the IAT, as a measure of individual differences or otherwise, indexes the same or different underlying latent constructs as parallel direct measures relying on self-report has been a matter of vigorous theoretical debate. Note that multiple well-established theoretical approaches (e.g., Cunningham & Zelazo, 2007; De Houwer, 2014; De Houwer et al., 2020; Fazio, 2007; Kruglanski & Gigerenzer, 2011; Van Bavel et al., 2012) are consistent with the possibility that, rather than direct measures tapping explicit constructs and indirect measures tapping implicit constructs, both types of measure may reflect largely overlapping mental content. From this perspective, IATs may very well be valid indicators of attitudes, beliefs, and self-concept that measure these constructs under conditions of automaticity even if IAT scores are highly correlated, or even fully redundant, with scores derived from direct measures.

Although some of its claims may appear to be more general, Schimmack's article focuses solely on the IAT as a measure of individual differences in implicit attitudes; as such, it does not speak to the bulk of the studies relying on the IAT. Such studies include experimental work (Cone et al., 2017; De Houwer et al., 2020; Gawronski & Bodenhausen, 2006); studies in which IATs are used to improve the prediction of behaviors that are of inherent practical interest, such as in work on suicide and nonsuicidal self-injury (Barnes et al., 2016; C. R. Glenn et al., 2019; J. J. Glenn et al., 2017); and a quickly growing set of investigations that uses IATs to study societal phenomena at the level of regions rather than individuals (for reviews, see Hehman, Calanchini, Flake, & Leitner, 2019; Payne et al., 2017a). Evidence of validity must be established for each new application of a measure; however, the MTMM approach does not seem appropriate in any of these contexts.

In summary, the evidence provided by Schimmack is perfectly consistent with the potential of the IAT to serve as a valid measure of attitudes, a valid measure

of automatic cognition, a valid measure of individual differences, and, ironically given his central claim, even a valid measure of individual differences in automatically revealed attitudes. Schimmack's analyses seem relatively difficult to reconcile with a view that posits a clean separation of processes driving responding on direct and indirect measures of social cognition. We believe that the debate between dual-process theories and their more recent alternatives may indeed benefit from stronger inclusion of an individual-difference perspective. However, the theoretical debate cannot be decided by relying exclusively on MTMM investigations. And, crucially, the validity of the IAT as a measure of automatically revealed associations does not hinge on any particular outcome of that debate.

Transparency

ORCID iD

Benedek Kurdi 🆔 https://orcid.org/0000-0001-5000-0584

Acknowledgments

Notes

1. To avoid any misunderstanding, we use the qualifiers *direct* and *indirect* to refer to measures and the qualifiers *explicit* and *implicit* to refer to underlying latent constructs. As we discuss in more detail later, it should be noted that the validity and utility of indirect measures, such as the Implicit Association Test, do not hinge on the existence of separate explicit and implicit representations in memory.

2. We use the term *association* as a shorthand to refer to any underlying mental representation that implies that one category (e.g., "planes") goes together relatively more with one attribute (e.g., "fast") and the other category (e.g., "trains") goes together relatively more with the other attribute (e.g., "slow"). This mental representation could be a set of associations (such as planes–fast and trains–slow), a set of propositions (e.g., planes are fast and trains are slow), or some other alternative. Moreover, we are also sympathetic to the usefulness of defining implicit evaluations as a behavior rather than a mental construct (De Houwer, 2019).

3. At the same time, although modeling decisions always represent a compromise, they should be supported by well-reasoned arguments. Readers interested in a more statistically oriented commentary on the strength of the analytic decisions made by Schimmack should refer to the response by Vianello and Bar-Anan (2021).

4. On p. 400, Schimmack (2021) cites Cronbach as arguing that independent researchers may be better positioned to examine the validity of a test than its originators. Although we cannot speak for the originators of the Implicit Association Test, as researchers who have used it in numerous studies, we found little that was objectionable or even surprising in Schimmack's arguments.

5. It should also be noted that social-cognition researchers have repeatedly warned against using the Implicit Association Test as a diagnostic tool given these very debates on the relative contributions of persons versus situations to responding. Kurdi et al. (2019) stated that "given such malleability, we have always advised against using a single intergroup IAT as a device for the selection of people, such as whether to hire someone for a job or admit them to a club. The measure is of value in two contexts: research and education" (p. 16). Likewise, the Project Implicit educational website notes the following:

> The IAT has potential for use beyond the laboratory; however, there are problems with using it outside of the safeguards of a research institution. First, people may use the IAT to make decisions about themselves (e.g., what should I buy, where should I go to school?). Second, people may use it to make decisions about others (e.g., does this potential job candidate have racial bias?). On the surface these might seem like acceptable uses; however, we assert that the IAT should not be used in any such ways. We cannot be certain that any given IAT can diagnose an individual. At this stage in its development, it is preferable to use the IAT mainly as an educational tool to develop awareness of implicit preferences and stereotypes. For example, using the IAT to choose jurors is not ethical. In contrast, it might be appropriate to use the IAT to teach jurors about the possibility of unintended bias. Using the IAT to make significant decisions about oneself or others could lead to undesired and unjustified consequences. (Project Implicit, 2019)

6. Of course, if all meaningful variance in attitudes were to be a function of the situation with the person accounting for no variability whatsoever, this finding would make an individual-difference approach to attitudes futile.

7. We use the term *automatic* (rather than *implicit*) here to highlight the point that the valid measurement of processes that unfold in a relatively unconscious, unintentional, efficient, and controlled manner does not presuppose the existence of implicit-memory representations, entirely dissociated from explicit ones, in memory. At the same time, automaticity refers not only to the operating conditions of measures but also to features of the underlying processes that they are intended to capture.

8. Most of our discussion here focuses on the construct of attitude (i.e., evaluations along a positive-negative continuum). However, the same arguments would also apply to Implicit Association

Tests measuring different constructs, including stereotypes and self-concept.

References

Arkes, H. R., & 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

Axt, J. R. (2018). The best way to measure explicit racial attitudes is to ask about them. *Social Psychological and Personality Science*, *9*, 896–906. doi:10.1177/1948550617728995

Bar-Anan, Y., & Vianello, M. (2018). A multi-method multi-trait test of the dual-attitude perspective. *Journal of Experimental Psychology: General*, *147*, 1264–1272. doi:10.1037/xge0000383

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

Barnes, S. M., Bahraini, N. H., Forster, J. E., Stearns-Yoder, K. A., Hostetter, T. A., . . . Nock, M. K. (2016). Moving beyond self-report: Implicit associations about death/life prospectively predict suicidal behavior among veterans. *Suicide and Life-threatening Behavior*, *47*, 67–77. doi:10.1111/sltb.12265

Barnes-Holmes, D., Waldron, D., Barnes-Holmes, Y., & Stewart, I. (2009). Testing the validity of the Implicit Relational Assessment Procedure and the Implicit Association Test: Measuring attitudes toward Dublin and country life in Ireland. *The Psychological Record*, *59*, 389–406. doi:10.1007/BF03395671

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

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*, 1061–1071. doi:10.1037/0033-295X.111.4.1061

Calanchini, J., Sherman, J. W., Klauer, K. C., & Lai, C. K. (2014). Attitudinal and non-attitudinal components of IAT performance. *Personality and Social Psychology Bulletin*, *40*, 1285–1296. doi:10.1177/0146167214540723

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin*, *56*, 81–105. doi:10.1037/h0046016

Carlston, D. E., & Skowronski, J. J. (1994). Savings in the relearning of trait information as evidence for spontaneous inference generation. *Journal of Personality and Social Psychology*, *66*, 840–856. doi:10.1037//0022-3514.66.5.840

Cone, J., Mann, T. C., & Ferguson, M. J. (2017). Changing our implicit minds: How, when, and why implicit evaluations can be rapidly revised. In J. M. Olson (Ed.), *Advances in experimental social psychology* (Vol. 56, pp. 131–199). San Diego, CA: Academic Press. doi:10.1016/bs.aesp.2017.03.001

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. doi:10.1037/0022-3514.89.4.469

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671–684. doi:10.1037/h0043943

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

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

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *Trends in Cognitive Sciences*, *11*, 97–104. doi:10.1016/j.tics.2006.12.005

Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative reprocessing model: A multilevel framework for attitudes and evaluation. *Social Cognition*, *25*, 736–760. doi:10.1521/soco.2007.25.5.736

De Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass*, *8*, 342–353. doi:10.1111/spc3.12111

De Houwer, J. (2019). Implicit bias is behavior: A functional-cognitive perspective on implicit bias. *Perspectives on Psychological Science*, *14*, 835–840. doi:10.1177/1745691619855638

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

De Houwer, J., Van Dessel, P., & Moran, T. (2020). Attitudes beyond associations: On the role of propositional representations in stimulus evaluation. In B. Gawronski (Ed.), *Advances in experimental social psychology* (Vol. 61, pp. 127–184). San Diego, CA: Academic Press. doi:10.1016/bs.aesp.2019.09.004

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich College Publishers.

Embretson, S. E. (1998). A cognitive design system approach to generating valid tests: Application to abstract reasoning. *Psychological Methods*, *3*, 380–396. doi:10.1037/1082-989X.3.3.380

Falk, C. F., Heine, S. J., Takemura, K., Zhang, C. X. J., & Hsu, C.-W. (2014). Are implicit self-esteem measures valid for assessing individual and cultural differences? *Journal of Personality*, *83*, 56–68. doi:10.1111/jopy.12082

Faure, R., Righetti, F., Seibel, M., & Hofmann, W. (2018). Speech is silver, nonverbal behavior is gold: How implicit partner evaluations affect dyadic interactions in close relationships. *Psychological Science*, *29*, 1731–1741. doi:10.1177/0956797618785899

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). San Diego, CA: Academic Press. doi:10.1016/S0065-2601(08)60318-4

Fazio, R. H. (2007). Attitudes as object–evaluation associations of varying strength. *Social Cognition*, *25*, 603–637. doi:10.1521/soco.2007.25.5.603

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238. doi:10.1037/0022-3514.50.2.229

Gawronski, B. (2019). Six lessons for a cogent science of implicit bias and its criticism. *Perspectives on Psychological Science*, *14*, 574–595. doi:10.1177/1745691619826015

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692–731. doi:10.1037/0033-2909.132.5.692

Gawronski, B., & Bodenhausen, G. V. (2011). The associative–propositional evaluation model: Theory, evidence, and open questions. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 44, pp. 59–127). San Diego, CA: Academic Press. doi:10.1016/B978-0-12-385522-0.00002-0

Gawronski, B., & Bodenhausen, G. V. (2017). Beyond persons and situations: An interactionist approach to understanding implicit bias. *Psychological Inquiry*, *28*, 268–272. doi:10.1080/1047840X.2017.1373546

Glenn, C. R., Millner, A. J., Esposito, E. C., Porter, A. C., & Nock, M. K. (2019). Implicit identification with death predicts suicidal thoughts and behaviors in adolescents. *Journal of Clinical Child & Adolescent Psychology*, *48*, 263–272. doi:10.1080/15374416.2018.1528548

Glenn, J. J., Werntz, A. J., Slama, S. J. K., Steinman, S. A., Teachman, B. A., & Nock, M. K. (2017). Suicide and self-injury-related implicit cognition: A large-scale examination and replication. *Journal of Abnormal Psychology*, *126*, 199–211. doi:10.1037/abn0000230

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., & Banaji, M. R. (2017). The implicit revolution: Reconceiving the relation between conscious and unconscious. *American Psychologist*, *72*, 861–871. doi:10.1037/amp0000238

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., 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

Hahn, A., & Gawronski, B. (2019). Facing one's implicit biases: From awareness to acknowledgment. *Journal of Personality and Social Psychology*, *116*, 769–794. doi:10.1037/pspi0000155

Hahn, A., Judd, C. M., Hirsh, H. K., & Blair, I. V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General*, *143*, 1369–1392. doi:10.1037/a0035028

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*, 1166–1186. doi:10.3758/s13428-017-0935-1

Hehman, E., Calanchini, J., Flake, J. K., & Leitner, J. B. (2019). Establishing construct validity evidence for regional measures of explicit and implicit racial bias. *Journal of Experimental Psychology: General*, *148*, 1022–1040. doi:10.1037/xge0000623

Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorizing in implicit attitude research: Propositional and behavioral alternatives. *The Psychological Record*, *61*, 465–496.

Irving, L. H., & Smith, C. T. (2020). Measure what you are trying to predict: Applying the correspondence principle to the Implicit Association Test. *Journal of Experimental Social Psychology*, *86*, Article 103898. doi:10.1016/j.jesp.2019.103898

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. doi:10.1037/0022-3514.93.3.353

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, *118*, 97–109. doi:10.1037/a0020762

Kurdi, B., & Banaji, M. R. (2017). Reports of the death of the individual difference approach to implicit social cognition may be greatly exaggerated: A commentary on Payne, Vuletich, and Lundberg. *Psychological Inquiry*, *28*, 281–287. doi:10.1080/1047840X.2017.1373555

Kurdi, B., Seitchik, A. E., Axt, J. R., Carroll, T. J., Karapetyan, A., Kaushik, N., . . . Banaji, M. R. (2019). Relationship between the Implicit Association Test and intergroup behavior: A meta-analysis. *American Psychologist*, *74*, 569–586. doi:10.1037/amp0000364

Lindgren, K. P., Neighbors, C., Teachman, B. A., Baldwin, S. A., Norris, J., Kaysen, D., . . . Wiers, R. W. (2016). Implicit alcohol associations, especially drinking identity, predict drinking over time. *Health Psychology*, *35*, 908–918. doi:10.1037/hea0000396

Lindgren, K. P., Westgate, E. C., Kilmer, J. R., Kaysen, D., & Teachman, B. A. (2012). Pick your poison: Stimuli selection in alcohol-related implicit measures. *Addictive Behaviors*, *37*, 990–993. doi:10.1016/j.addbeh.2012.03.024

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, *3*, 635–694. doi:10.2466/pr0.1957.3.3.635

McConnell, A. R., & Rydell, R. J. (2014). The systems of evaluation model: A dual-systems approach to attitudes. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 204–217). New York, NY: Guilford Press.

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. doi:10 .1037/a0030734

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*, 741–749. doi:10.1037/0003-066X.50.9.741

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234. doi:10.1037/h0031564

Mischel, W. (1968). *Personality and assessment*. New York, NY: Wiley.

Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory & Cognition*, *4*, 648–654. doi:10.3758/ BF03213230

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*, 565–584. doi:10.1037/0096-3445 .134.4.565

Nosek, B. A. (2007). Implicit–explicit relations. *Current Directions in Psychological Science*, *16*, 65–69. doi:10 .1111/j.1467-8721.2007.00477.x

Nuttin, J. M. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology*, *15*, 353–361. doi:10.1002/ejsp.2420150309

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, *105*, 171–192. doi:10 .1037/a0032734

Payne, B. K., Vuletich, H. A., & Lundberg, K. B. (2017a). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, *28*, 233–248. doi:10.1080/1047840X.2017.1335568

Payne, B. K., Vuletich, H. A., & Lundberg, K. B. (2017b). Flipping the script on implicit bias research with the bias of crowds. *Psychological Inquiry*, *28*, 306–311. doi:10.10 80/1047840X.2017.1380460

Project Implicit. (2019, December 2). Ethical considerations. Retrieved from https://implicit.harvard.edu/implicit/eth ics.html

Rae, J. R., & Greenwald, A. G. (2017). Persons or situations? Individual differences explain variance in aggregated implicit race attitudes. *Psychological Inquiry*, *28*, 297–300. doi:10.1080/1047840X.2017.1373548

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*, 995–1008. doi:10.1037/0022-3514.91.6.995

Schimmack, U. (2021). The Implicit Association Test: A method in search of a construct. *Perspectives on Psychological Science*, *16*(2), 396–414. doi:10.1177/1745691619863798

Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*, *25*, 638–656. doi:10.1521/ soco.2007.25.5.638

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108–131. doi:10.1207/ S15327957PSPR0402_01

Sternberg, R. J. (1981). Testing and cognitive psychology. *American Psychologist*, *36*, 1181–1189. doi:10.1037/0003-066X.36.10.1181

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

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

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*, 440–463. doi:10.1037/a0018963

Van Bavel, J. J., Xiao, Y. J., & Cunningham, W. A. (2012). Evaluation is a dynamic process: Moving beyond dual system models. *Social and Personality Psychology Compass*, *6*, 438–454. doi:10.1111/j.1751-9004.2012.00438.x

Van Dessel, P., Gawronski, B., & De Houwer, J. (2019). Does explaining social behavior require multiple memory systems? *Trends in Cognitive Sciences*, *23*, 368–369. doi:10.1016/j.tics.2019.02.001

van Harmelen, A.-L., De Jong, P. J., Glashouwer, K. A., Spinhoven, P., Penninx, B., & Elzinga, B. M. (2010). Child abuse and negative explicit and automatic self-associations: The cognitive scars of emotional maltreatment. *Behaviour Research and Therapy*, *48*, 486–494. doi:10.1016/j.brat.2010.02.003

Vianello, M., & Bar-Anan, Y. (2021). Can the Implicit Association Test measure automatic judgment? The validation continues. *Perspectives on Psychological Science*, *16*(2), 415–421. doi:10.1177/1745691619897960

Whitely, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, *93*, 179–197. doi:10.1037/0033-2909.93.1.179