Commentary

Being on the Lookout for Validity

Comment on Sriram and Greenwald (2009)

Malte Friese[1] and Klaus Fiedler[2]

[1]University of Basel, Switzerland
[2]University of Heidelberg, Germany

Abstract. Recently, Sriram and Greenwald (2009) introduced a new IAT-like measure, the Brief Implicit Association Test (BIAT). Because the BIAT is a new development, empirical evidence for its validity is yet scarce. This comment focuses on two possible approaches to validation research on the BIAT: (1) a pragmatic correlational approach and (2) an experimental approach aiming at causal understanding of the BIAT task. We argue that both approaches provide valuable and mutually complementing evidence, but only experimental research can conclusively show that the to-be-measured constructs causally influence BIAT scores. Because such a causal analysis is at the core of the validity problem, research on the BIAT should reduce the asymmetry in favor of correlational validation that emerged in traditional IAT research.

Keywords: implicit measures, Implicit Association Test, validity

Although the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) has been extensively studied and applied in various subdisciplines of psychology and beyond, its validity remains an intensively debated issue (e.g., Blanton et al., 2009; Fiedler, Messner, & Bluemke, 2006; Greenwald, Poehlman, Uhlmann, & Banaji, 2009). Recently, Sriram and Greenwald (2009; hereafter S&G) have introduced a new variant of an IAT-like measure, the Brief Implicit Association Test (BIAT). Like an IAT, the BIAT involves two target categories (*Coke* vs. *Pepsi*) and two attributes (*pleasant* vs. *unpleasant*), but participants are instructed to focus on just two of the four categories in each of two blocks, using one response key for the focal categories and another response key for nonfocal categories. For example, *Coke* and *pleasant* may be focal in one block and *Pepsi* and *pleasant* in the other block. Faster response latencies in the former block would be expected when the attitude toward Coke is more positive than that toward Pepsi.

Several assets of the BIAT have to be acknowledged. First, its administration time is considerably faster than that for the traditional IAT, allowing for several BIATs in a single session. Second, internal consistencies were encouraging given the high number of BIATs in some studies. Third, although original IAT scores may be more difficult to control than self-report measures, experience with the IAT enables individuals to influence their scores (e.g., Fiedler & Bluemke, 2005; Steffens, 2004). The low number of BIAT trials may serve to impede strategic and volitional influences and thereby control for unwanted simulation and dissimulation effects.

Because the BIAT is a new development, the evidence for its validity is scarce to date. In the remainder of this comment, we will juxtapose two possible validation strategies for the BIAT, a pragmatic correlational approach and an experimental approach aiming at a causal understanding of the BIAT task. Both strategies can lead to valuable, mutually complementing evidence, even though only the experimental approach may reveal unambiguous answers. Although we discuss these issues in relation to the BIAT, they also apply to other measures of psychological constructs more generally (cf. De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009).

Correlational Validation

The correlational approach aims at investigating the relations of psychological constructs. Drawing on theoretical expectations, researchers make predictions about the direction and size of these relations. They test their hypotheses by correlating the scores of the focal measure with criterion measures of (un-)related constructs or behavioral outcomes. Ideally, as more and more branches of the nomological network (Cronbach & Meehl, 1955) are supported by empirical evidence, correlational validation narrows down the constructs that can possibly account for the empirical pattern of results, thereby strengthening the confidence that the measure under investigation indeed reflects the construct it is intended to measure.

A host of studies have employed this approach to validate the traditional IAT as a measure of implicit attitudes, stereotypes, or self-concepts (for reviews, see Friese, Hofmann, & Schmitt, 2008; Greenwald et al., 2009; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Hofmann, Gschwendner, Nosek, & Schmitt, 2005; Nosek, 2005; Nosek et al., 2007). S&G relied on correlations of the BIAT with parallel self-report measures as the only validity criterion in many cases. We deem this confinement to implicit-explicit correlations unsatisfactory as it investigates only a minimal part of the nomological network. Future

DOI: 10.1027/1618-3169/a000051

studies should spell out the network defining the BIAT's construct validity more completely, going beyond introspective self-report measures of the same construct, and they should provide more cogent evidence. Of special interest are correlations of the BIAT with manifest behaviors, considered the gold standard of correlational validation by many (e.g., Nosek & Greenwald, 2009). From a practical perspective in this matter it is not only important to know which behaviors the BIAT predicts, but also to demonstrate (a) when the BIAT predicts these behaviors and when not; (b) the BIAT's incremental validity compared to parallel self-report measures; and (c) its advantages over the traditional IAT. None of these issues have yet been addressed by S&G, leaving fields of interesting research questions to be answered.

Experimental Validation

Even the most comprehensive correlational validation, though, provides only partial satisfaction because it does not allow for causal inferences about the origins of the test results (De Houwer et al., 2009; Fiedler et al., 2006; Nosek & Smyth, 2007). To show that a measurement outcome is causally produced by the construct of interest is, however, the core of the concept of validity (Borsboom, Mellenbergh, & van Heerden, 2004). Correlations can be misleading, for example, because an unknown third variable may be responsible for an observed correlation pattern. For a compelling illustration, think of the phenomenon of Clever Hans (Pfungst, 1911), the fascinating horse that could seemingly solve arithmetic problems, communicating the numerical solution by knocking his foreleg an appropriate number of times on the ground. Although Hans' test results were highly correlated with the arithmetic criterion, this was in fact not a valid indicator of the horse's ability to calculate. What made Hans clever actually was not calculation but utilization of a strategy, which consisted of attending to the tester's subtle nonverbal signals.[1] Hans had learned to truncate counting with his foreleg as soon as the tester sent out unintended cues that the correct solution was reached. Because the tester knew the correct answer, the nonverbal message coincided with the correct solution, so that Hans' strategy mimicked accurate calculation.

If the goal is only to predict the correct solution, does it matter how Hans finds it, through calculation or nonverbal cues? – Indeed, it does matter, because the correlation that "validates" Hans' performance is labile and context specific. When the specific test conditions change (e.g., nonverbal cues eliminated), the noncausal correlation disappears. Returning to the BIAT, by analogy, it also matters whether its correlation with verbal or behavioral criteria is causal or only incidental or spurious. Do faster latencies when focusing on *Coke* and *pleasant* than on *Pepsi* and *pleasant* really reflect an attitude or a strategy that makes BIAT life easier, analogous to Hans' strategy? Could the correlation of the BIAT latency difference with self-reported preferences reflect a strategy rather than a genuine attitude, defined as an association between *Coke* and *pleasant* in long-term memory?

Speeded-classification tasks are particularly prone to strategic influences that can make the respondents' life easier (cf. Fiedler et al., 2006; Garner, 1974; Rothermund & Wentura, 2004). It is therefore essential to rule out strategies that mimic attitudes or stereotypes. One simple, generic strategy that produces spurious IAT effects was recently outlined by Bluemke and Fiedler (in press). Rather than explaining this strategy at length, let us apply it to the Hans' example. Engage in the following thought experiment. Imagine Clever Hans performs a Coke-Pepsi BIAT. Imagine further that Hans does prefer Coke over Pepsi, as readily expressed in self-report. His significant BIAT score (favoring Coke) might nevertheless partly reflect a strategy, rather than an association between *Coke* and *pleasant*, simply because of a twofold response bias. Hans may be (a) more ready to respond pleasant than unpleasant (a common finding in evaluative priming) and (b) more ready to respond Coke than Pepsi (either because Hans prefers Coke or because Coke is more prominent). If both prevailing biases call for the same response (i.e., in the Coke-pleasant focus condition), this should of course be easier than when there is a conflict between two biases pointing to different response keys.

Does it matter? – The answer is: It depends. If Hans' strategic bias toward Coke reflects his true preference, then the strategy predicts Hans' attitude. The more positive Hans' attitude toward Coke, the stronger his bias and the stronger his BIAT score due to a double response bias. However, a causal analysis may reveal that the bias toward Coke is independent of Hans' attitude. It may simply reflect that Coke is more popular or prevalent in advertising, even to Hans' friend Otto, who likes Coke and Pepsi equally well. Otto may show the same double bias, resulting in the same BIAT score as Hans, though for different reasons. The same double response strategy reflects an attitude in Hans but not in Otto.

Note that such a strategic account can explain that if participants actually differ in their (honestly reported) attitude toward Coke, the BIAT effect should correlate with the actual attitude. After all, an attitude is a sufficient cause for a response bias. However, it is not a necessary cause because biases can reflect many other factors (e.g., Coke's popularity). Thus, hypothetically assuming a *single* causal influence on BIAT scores (e.g., attitude or some other factor), the reported correlation may validate Hans' BIAT score but not Otto's. Assuming *manifold* causal influences, those that are unrelated to the attitude may decrease validity and correlations. For some individuals the attitude may contribute more strongly to a BIAT score than for others. In the absence of cogent causal evidence, correlations remain equivocal as an indicator of attitudes.

[1] Note that the term "strategy" in this context is not meant to (necessarily) imply conscious, intentional efforts, but rather in the sense it is often used in signal-detection frameworks, where strategies are frequently described as unconscious inductions of response biases (Brown, Steyvers, & Hemmer, 2007; Wixted, 1993).

The BIAT's focus instruction may actually be an effective way of inducing response biases. Response facilitation can be expected when the instruction focus is compatible with other sources of biases. Assuming a bias toward *Kerry* and against *Bush* and granting a general bias toward *good* (rather than *bad*), explicit instructions to focus on *Kerry/good* will be compatible with existing biases and lead thereby to faster responding than instructions to focus on *Bush/good*. Again, because the bias toward *Kerry* can be assumed to increase with an increasing attitude toward *Kerry*, the facilitation effect may correlate with actually existing differences in attitude (regardless of other possible determinants of a Kerry bias).

However, when the constant focus on *good* is replaced by *bad*, the instruction to focus on *Kerry/bad* will no longer be compatible with a joint bias on *Kerry* and *good*. Thus, an analysis of strategic response biases suggests a plausible account for the divergent findings reported by S&G for different constant-focus conditions. When a constant focus on *good* and *self* is consistent with affective biases, the BIAT is reliable and clearly correlated with actual preferences. In contrast, when a constant focus on bad and others is incompatible with existing affective biases, the BIAT drops sharply in reliability and convergent validity.

We cannot know, of course, whether this tentative account is correct. However, in any case, validation research should devote more thought to a causal analysis of the task beyond correlational analyses to arrive at a deeper understanding of BIAT scores (cf. De Houwer et al., 2009). With this conclusion in mind, we will finally discuss three open questions with regard to the BIAT: (1) Does the focus on just two categories reduce "spontaneous variation in subject strategies?" (p. 283); (2) Does the BIAT measure implicit attitudes and stereotypes? (p. 283); and (3) Are there other causal factors that obscure the BIAT measure?

Regarding the first question, more consistent subject strategies should reduce error variance in BIAT scores. This reduced error variance should in turn increase the systematic variance, leading to enhanced validity. Surprisingly, S&G did not test this supposed advantage over the traditional IAT. To this end, data is needed that sheds light on the strategies that participants use to master the BIAT with and without focus instructions. Even more crucially, it needs to be shown that the focus instructions indeed boost the systematic variance compared to BIATs without the focus instruction (as manifested in more substantial correlational patterns and a stronger impact of the to-be-measured construct on BIAT scores). In a related vein, if the focus instructions increase the validity of the BIAT they may also increase the validity of the traditional IAT. This appears like an attractive research question to investigate.

Second, if the BIAT is to measure implicit attitudes and stereotypes, these constructs need to *causally* influence BIAT scores. To investigate this assumption, this first demands a thorough definition of these constructs, a step that in itself is highly disputed (e.g., Fazio & Olson, 2003; Gawronski, Hofmann, & Wilbur, 2006; Greenwald & Banaji, 1995; Wilson, Lindsey, & Schooler, 2000). Furthermore, experimental manipulations aimed at (temporarily) altering the construct should be reflected in varying BIAT scores (De Houwer et al., 2009).

Third, we have little doubt that variations in attitudes and stereotypes can cause variation in BIAT scores. However, theory and research on the traditional IAT suggest that numerous other factors can influence BIAT scores as well. This leaves researchers in doubt about the diagnostic meaning of IAT scores. Unwanted factors that causally contribute to the magnitude of IAT effects include, among others, the order of combined blocks (Nosek, Greenwald, & Banaji, 2006), the composition of the stimulus sets (e.g., Bluemke & Friese, 2006; Govan & Williams, 2004), salience asymmetries (Rothermund & Wentura, 2001, 2004; see also Proctor & Cho, 2006), task-switching costs (Klauer & Mierke, 2005; Mierke & Klauer, 2001), the social context (Dasgupta & Greenwald, 2001; Lowery, Hardin, & Sinclair, 2001), or the application of mental strategies in mastering the IAT task (e.g., Blair, Ma, & Lenton, 2001; Fiedler & Bluemke, 2005; Steffens, 2004).

These findings suggest that BIAT scores may reflect a complex mix of causal influences, of which attitudes, stereotypes, and self-concept are only a part. They highlight the need for a testable process model to experimentally disentangle and estimate the various sources of variance for a given BIAT score. Such a model would enhance the understanding of how BIAT and IAT scores translate into the constructs of interest (Fiedler et al., 2006; Nosek & Smyth, 2007).

At the same time, experimental research also has its limitations and pitfalls. More specifically, the aspired evidence for the causal role of a certain attribute on BIAT scores presupposes that the employed experimental manipulation indeed affects the intended attribute. This requires the attribute to be malleable to a certain extent. In addition, researchers need to make sure that their experimental manipulations affect the target attribute, but only the target attribute and not also unintentionally some other attribute that is capable of causally influencing BIAT scores (Deutsch & Gawronski, 2009). These premises are difficult to meet in some cases, demanding profound theoretical and methodological expertise (see De Houwer et al., 2009).

Conclusions

In this comment we have discussed two approaches to validation of the BIAT, correlational and causal-experimental validation. Correlational validation is largely mute to the question of whether the to-be-measured construct has actually caused the measurement outcome, in this case BIAT scores. Experimental research allows for causal inferences and is therefore suited to fill this gap. Although one never knows whether the experimentally controlled BIAT factors actually tap on the determinants of BIAT scores outside the laboratory, such experimental evidence is valuable to set apart causal from spurious correlations.

To the extent that a test is contaminated with unwanted extraneous factors, its validity is cast into doubt. Consequently, the numerous factors that have been found to affect traditional IAT scores are a threat to the BIAT's validity as well. They threaten the diagnostic value of individual BIAT scores and mean group effects and open the door to false diagnostic conclusions. On the positive side, there has been a multitude of studies on the traditional IAT that found theory-consistent correlations with a host of other measures and behaviors (e.g., Friese et al., 2008; Greenwald et al., 2009; Hofmann, Gawronski et al., 2005; Hofmann, Gschwendner et al., 2005; Nosek, 2005; Nosek et al., 2006, 2007). Often, these correlations emerged as a function of a theoretically derived third variable such as processing resources. It is important to note that these correlations emerged *despite* various undesired factors threatening the measures' validity, not due to these factors. Undesired influences such as the order of combined blocks or salience asymmetries could only account for these findings if one would assume that these factors were correlated with the to-be-measured construct or predicted the criterion variables for some other reason. If this had actually been the case, the validity of the IATs would ironically have been negligible in the sense that the to-be-measured constructs would not have causally influenced the measurement outcomes. However, the measures' pragmatic value (as opposed to causal validity) for predicting a range of behaviors would remain unaffected (Often times behaviors that parallel self-report measures did not predict as well as the IAT).

To conclude, both strategies, correlational and experimental validation, have their strengths, and both can provide valuable information about the validity of a measure. We feel that it is not an either/or decision between these two approaches in the validation of the BIAT. Both should receive researcher's attention. If we could make a wish, experimental validation of the BIAT should receive a greater weight to reduce the asymmetry in favor of correlational validation that emerged in research on the traditional IAT. Researchers should look out for converging evidence surfacing from different research strategies that speak to the validity of the BIAT.

Acknowledgments

We thank Matthias Bluemke for helpful comments on an earlier version of this manuscript.

References

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81*, 828–841.

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.

Bluemke, M., & Fiedler, K. (in press). Base rate effects on the IAT. *Consciousness and Cognition*.

Bluemke, M., & Friese, M. (2006). Do features of stimuli influence IAT effects?. *Journal of Experimental Social Psychology, 42*, 163–176.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review, 111*, 1061–1071.

Brown, S., Steyvers, M., & Hemmer, P. (2007). Modeling experimentally induced strategy shifts. *Psychological Science, 18*, 40–45.

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

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81*, 800–814.

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

Deutsch, R., & Gawronski, B. (2009). When the method makes a difference: Antagonistic effects on "automatic evaluations" as a function of task characteristics of the measure. *Journal of Experimental Social Psychology, 45*, 101–114.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54*, 297–327.

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Test. *Basic and Applied Social Psychology, 27*, 307–316.

Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the "I", the "A", and the "T": A logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology, 17*, 74–147.

Friese, M., Hofmann, W., & Schmitt, M. (2008). When and why do implicit measures predict behaviour? Empirical evidence for the moderating role of opportunity, motivation, and process reliance. *European Review of Social Psychology, 19*, 285–338.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

Gawronski, B., Hofmann, W., & Wilbur, C. J. (2006). Are "implicit" attitudes unconscious?. *Consciousness and Cognition, 15*, 485–499.

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

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.

Govan, C. L., & Williams, K. D. (2004). Reversing or eliminating IAT effects by changing the affective valence of the stimulus items. *Journal of Experimental Social Psychology, 40*, 357–365.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*, 1369–1385.

Hofmann, W., Gschwendner, T., Nosek, B. A., & Schmitt, M. (2005). What moderates implicit-explicit consistency? *European Review of Social Psychology, 16*, 335–390.

Klauer, K. C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence

for a task-set switching account of the Implicit Association Test effect. *Personality and Social Psychology Bulletin, 31*, 208–217.
Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81*, 842–855.
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.
Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134*, 565–584.
Nosek, B. A., & Greenwald, A. G. (2009). (Part of) the case for a pragmatic approach to validity: Comment on De Houwer, Teige-Mocigemba, Spruyt, and Moors (2009). *Psychological Bulletin, 135*, 373–376.
Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2006). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). New York, NY: Psychology Press.
Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the Implicit Association Test. Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology, 54*, 14–29.
Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., et al. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology, 18*, 36–88.
Pfungst, O. (1911). *Clever Hans*. New York: Holt.
Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin, 132*, 416–442.
Rothermund, K., & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test. *Zeitschrift für Experimentelle Psychologie, 48*, 94–106.
Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test (IAT): Dissociating salience from associations. *Journal of Experimental Psychology: General, 133*, 139–165.
Sriram, N., & Greenwald, A. G. (2009). The Brief Implicit Association Test. *Experimental Psychology, 56*, 283–294.
Steffens, M. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology, 51*, 165–179.
Wilson, T., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107*, 101–126.
Wixted, J. (1993). A signal detection analysis of memory for nonoccurrence in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 19*, 400–411.

Received June 12, 2009
Revision received July 28, 2009
Accepted July 29, 2009
Published online: December 23, 2009

Malte Friese

Department of Psychology
University of Basel
Missionsstr. 60/62
4055 Basel
Switzerland
Tel. +41 61 267 06 25
Fax +41 61 267 06 28
E-mail malte.friese@unibas.ch

