

Implicit Measures of Associations

A Case of Exaggerated Promises?

Anne Roefs, Jorg Huijding, Fren T. Y. Smulders,
Anita T. M. Jansen, and Colin M. MacLeod

Researchers in psychology have long sought ways to more fully understand human cognition and behavior that do not rely on self-report. Investigators in the domain of psychopathology are no exception. Three reasons are most often cited, harking back to the early critiques of introspection (see Boring, 1953): (1) Self-reports rely on introspection, yet we know that people do not have introspective access to all of the mechanisms that underlie their behavior (Nisbett & Wilson, 1977); (2) self-report measures are subject to socially desirable biases; and (3) people may dismiss as irrelevant certain cognitions that actually are relevant (Schwarz, 1999; Schwarz & Oyserman, 2001).

MacLeod (1993) even argued that accepting introspective self-report measures as data would place psychopathology research “outside the boundaries of legitimate science” (p. 179). He convincingly reasoned that the cognitive revolution in clinical psychology took quite a different form than the cognitive revolution in experimental psychology did. In experimental psychology, cognition was allowed in the form of hypothetical constructs in theories, but self-reported cognitions were not accepted as data. The evaluation of theories relied only on directly observable measures—that is, overt behaviors (e.g., response latencies). In clinical psychology, cognitions were allowed in theory as well as in data. So self-reports of cognition were accepted as data for testing

cognitive models of psychopathology. In a sense, this meant a return to the era before behaviorism. However, McNally (2001) has rightfully pointed out that it is neither possible nor right to dismiss introspection altogether, while also noting that introspection is not entirely absent in cognitive experimental psychology either, even when conducted in a neuroscience framework. Certain phenomena have no behavioral manifestation, are only accessible by introspection, and can only be reflected in verbal report (e.g., obsessions). He argues that researchers should be attentive to when introspection is useful and when it is not.

In psychopathology, research has focused for several decades on aspects of cognition such as biases in attention (e.g., Williams, Mathews, & MacLeod, 1996) and memory (e.g., Mitte, 2008), the goal being to better understand the dysfunctional cognitive architecture of patients. In the past couple of decades specifically, a different class of measurement procedures has gained popularity in psychopathology research. The most visible examples include the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) and the affective priming paradigm (APP; Fazio, Sanbonmatsu, Powell, & Kardes, 1986). These measurement procedures have been designed to obtain implicit measures of psychological attributes (for a review, see Roefs et al., 2011). The goal has been to develop measures untainted by the flaws of introspection and self-report.

What makes a measure implicit? De Houwer, Teige-Mocigemba, Spruyt, and Moors (2009) have defined an “implicit” measure as “a measurement *outcome* that is causally produced by the to-be-measured attribute in the absence of certain goals, awareness, substantial cognitive resources, or substantial time” (p. 350; emphasis in original). It is, however, important to keep in mind that implicitness is not an all-or-none feature of a measure (Moors & De Houwer, 2006). For example, a measure can be labeled as implicit in the sense that people are unaware of the existence of their association, or of its origin, or cannot control the process that leads to the measurement outcome. It is likely that not all implicit measures possess the same features of implicitness, and the degree to which a measure is implicit is to a large extent an empirical question. De Houwer and colleagues (De Houwer, 2006, 2009; De Houwer et al., 2009) provide extensive reviews of the degree to which various implicit measures meet the criteria for implicitness.

MEASUREMENT PROCEDURES

Quite a variety of measurement procedures have been used in the realm of psychopathology to obtain implicit measures of associations, with the

IAT (including variants on this procedure; Greenwald et al., 1998) and to a lesser extent the APP (Fazio et al., 1986) being the main ones (see Roefs et al., 2011). These paradigms are considered *indirect* measurement procedures. That is, participants are not asked to report directly on their thoughts or attitudes; instead, these are inferred from a pattern of response latencies. The *measurement outcomes* of these paradigms are considered *implicit measures* of associations. It is informative to try one of these procedures: IATs for a large number of topics can be found online (<https://implicit.harvard.edu/implicit>). Having taken the time to perform an IAT, it is then reasonable to consider these questions: (1) "Were you aware of what was being assessed?", (2) "Were you aware that you held the associations that the test identified?", and (3) "Do you agree that these associations reflect your attitude?" After taking a test once and reviewing your results, it would also be informative to try to "fake" the results by trying to appear different on the task the second time than you did the first time.

In this type of paradigm, participants are usually presented with single stimuli representing disorder-relevant targets (e.g., "spider") and attributes (e.g., "afraid"), and are instructed to respond as quickly as possible to these stimuli. As is common in cognitive measures, response latency and accuracy are the main dependent variables. The general idea behind these paradigms is that a certain pattern of response latencies and error percentages can serve as an index of the strength of target-attribute associations. Because they are the two most frequently used paradigms, the APP and the IAT are now explained in some detail.

In the APP (Fazio et al., 1986), two stimuli are presented in quick succession: a "prime" followed by a "target" (a stimulus onset asynchrony of 150 ms is typical). Prime stimuli represent the concepts of interest (e.g., spiders vs. butterflies), and target stimuli represent the attributes (e.g., positive vs. negative). A participant can ignore the prime, but must respond to the target as quickly as possible either by evaluating it as positive or negative or by naming it. The purpose of this paradigm is to assess the extent to which the prime presentation influences the response to the target. The critical idea is that affectively congruent prime-target pairs (e.g., "spider-awful") should lead to shorter response latencies than should affectively incongruent prime-target pairs (e.g., "spider-happy"). Mismatches should slow down responses. To the degree that such a pattern of response latencies is observed, it is taken to reflect the person's evaluation of the prime (e.g., spiders).

In the IAT (Greenwald et al., 1998), the participant is instructed to categorize each presented stimulus as quickly and as accurately as possible according to a target dimension (e.g., high-fat vs. low-fat foods) or an attribute dimension (e.g., positive vs. negative). In the two critical

combination phases of the IAT, the participant performs a double-categorization task; that is, he or she is instructed to make a binary decision for stimuli from two dimensions, with the dimension that is to be used to categorize the stimulus alternating on each trial. The two combination phases differ in the types of attribute and target stimuli that are mapped onto one response. So, for example, in a first combination phase, the participant presses the left button for high-fat foods and positive words, and the right button for low-fat foods and negative words. In a second combination phase, the response assignment for the target dimension is reversed (e.g., press the left button for low-fat foods and positive words; press the right button for high-fat foods and negative words). The IAT effect is typically computed as the difference in average response latency (and/or percentage of errors) between these two combination phases (see Greenwald, Nosek, & Banaji, 2003, for an alternative scoring algorithm, the *D*-measure). The logic behind the IAT is that people perform better when two associated targets/attributes share a response key than when two unassociated targets/attributes share a response key. So, in the current example, if participants are better at responding with the combination "high-fat and positive" versus "low-fat and negative" than when the combination is reversed, the conclusion will be that participants have a more positive association with high-fat foods than with low-fat foods.

COGNITIVE THEORY

In psychopathology research, an implicit measure is often assumed to reflect the strength of association between a disorder-relevant target (e.g., the self in depression) and an attribute (e.g., negative). Thus, for example, when depressed an individual would see him- or herself in a more negative light. These targets and attributes are usually derived from the proposed dysfunctional beliefs held by those with the disorder. These dysfunctional beliefs are, on the one hand, often apparent from the characterization of the disorder as specified in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association, 2013); on the other hand, they are specified in the principal cognitive theories, often based on Beck's cognitive approach to psychopathology (e.g., Beck & D. A. Clark, 1997; Beck, Freeman, & Davis, 2004; Beck, Rush, Shaw, & Emery, 1979).

For some disorders, not surprisingly, multiple competing cognitive models exist, and these are often more specific than Beck's schema model. Examples include D. M. Clark and Wells's (1995) model of social phobia; the impaired-cognitive-control account of depression (Joormann, Yoon,

& Zetsche, 2007); Ingram's (1984) information-processing analysis of depression; and the framework of Williams, Watts, MacLeod, and Matthews (1997) for depression. Similarly, Segal (1988) argues that for proving the existence and crucial role of negative self-schemas in depression, it is important to move beyond a heuristic use of the term "schemas" and a reliance of self-reports to test them. The conceptualization and testing of schemas should fit with current theorizing and methodologies of cognitive science.

In fact, however, much of the cognitive psychological research into the bases of psychopathology does not test more specific models and is (implicitly) based on Beck's schema model, which does not really address issues relating to the exact structure and function of schemas (Segal, 1988). In essence, the schema model proposes that a patient has a dysfunctional schema—that is, a cognitive assembly of associated knowledge that guides information processing in both automatic and controlled ways (see, e.g., Bartlett, 1932; Rumelhart, 1980). Therefore, information processing is thought to be biased, as reflected in measurable attention, memory, and interpretation biases in patient populations that differ from those of the general population. Implicit measures of associations were designed—and are believed—to have the capacity to reveal aspects of these dysfunctional schemas that explicit measures cannot reveal.

IMPLICIT MEASURES OF ASSOCIATION IN PSYCHOPATHOLOGY RESEARCH

We recently reviewed studies that obtained implicit measures of associations in various forms of psychopathology (Roefs et al., 2011). In the following paragraphs, then, rather than repeat this extensive summary, we limit ourselves to the general conclusions and most striking results from this review. The first thing to realize is that a large number of studies using implicit measures have provided evidence converging with that obtained with explicit measures. That is, in this set of studies, both types of measures have pointed in the same direction. For example, in patients suffering from spider phobia, implicit measures have generally indicated negative associations with spiders (e.g., Huijding & de Jong, 2007). In another set of studies, quite consistent evidence has been marshaled for implicit measures' going in the opposite direction of explicit measures. For example, in people suffering from depressive disorder, implicit measures quite consistently have reflected positive associations with the self (e.g., De Raedt, Schacht, Franck, & De Houwer, 2006); this is not consistent with the clinical picture and cognitive theory (e.g., D. A. Clark,

Beck, & Alford, 1999), where negative self-esteem is typically postulated. In a third set, implicit measures of associations with craved substances (food, alcohol, nicotine, and drugs) have reflected both the positive (e.g., "High-fat foods taste good") and the negative (e.g., "Alcohol can give you a hangover") aspects of these substances (e.g., Houben & Wiers, 2006; Roefs & Jansen, 2002). The picture is a complex one.

Consider the case of alcohol. Although both positive and negative associations with alcohol have been observed, multiple-regression analyses showed that only the positive associations proved to be predictive of drinking behavior (e.g., Houben & Wiers, 2008). In other disorders, implicit measures have shown incremental predictive validity, explaining a range of behavioral measures beyond those explained by explicit measures. This was the case, for example, with performance on a spider behavioral avoidance test (e.g., Teachman & Woody, 2003), panic symptoms (e.g., Teachman, Smith-Janik, & Saporito, 2007), mirror avoidance (Clerkin & Teachman, 2009), and food choice (e.g., Perugini, 2005). Thus, to take one example in a little more detail, an IAT measure of spider associations predicted variance of performance on a spider behavioral avoidance test, beyond what was explained by self-reported spider fear (e.g., Teachman & Woody, 2003).

In addition, in keeping with dual-process models of information processing (e.g., Fazio & Towles-Schwen, 1999), which typically posit two types of cognitive processes (automatic vs. controlled), some evidence has been found for the idea that implicit measures are specifically predictive of *spontaneous* types of behavior. The argument is that when more elaborative processing is possible, then behavior will be more in line with explicit measures. People then have the possibility to consciously consider both their cognitions and their behavior. Examples include the startle response in the context of specific phobia (Huijding & de Jong, 2006) and immediate affective reactions in the context of depression (Haefffel et al., 2007). Also in keeping with dual-process models are the findings that implicit measures have been shown to be predictive of behavior only when cognitive resources were limited either by an experimental manipulation (such as a depletion manipulation; e.g., Hofmann, Rauch, & Gawronski, 2007) or by an individual differences factor (e.g., Hofmann, Friese, & Roefs, 2009). This evidence for incremental predictive power demonstrates the added value that implicit measures can provide.

Though the experimental approach—that is, *manipulating* the theorized determinants of implicit measures—may be especially revealing, relatively few studies have taken this approach. Those that have done so generally have found the expected effect of the manipulation on the implicit measure(s). For example, a negative mood induction led to reduced self-positive associations in formerly depressed participants

(e.g., Gamar, Segal, Sagrati, & Kennedy, 2001), and inductions of craving and hunger positively affected associations with food in both healthy and eating-disordered participants (e.g., Seibt, Häfner, & Deutsch, 2007). On the one hand, this is evidence for the construct validity of the implicit measure; on the other hand, it is potentially problematic for the assessment of presumably stable dysfunctional schemas that these measures are so readily influenced by “acute” experimental manipulations.

Finally, a few studies have addressed the effect of treatment on disorder-relevant implicit measures, an area that surely warrants more attention. For specific phobia, social phobia, pain disorder, and panic disorder, successful exposure therapy or cognitive-behavioral therapy (CBT) was associated with significant improvements on implicit measures of disorder-relevant associations (e.g., Grumm, Erbe, von Collani, & Nestler, 2008; Teachman & Woody, 2003). For obesity and alcohol dependence, no or very small effects of treatment on implicit measures were observed (Craeynest, Crombez, Deforche, Tanghe, & De Bourdeaudhuij, 2008; Thush et al., 2009; Wiers, van de Luitgaarden, van den Wildenberg, & Smulders, 2005). Interestingly, for panic disorder, cognitive changes preceded symptom reduction (Teachman, Marker, & Smith-Janik, 2008)—suggesting, but not proving with certainty, that the cognitive changes may actually be the cause of the clinical improvement. The results of this latter study notwithstanding, an important question that remains for further research is whether observed changes in implicit measures after treatment are causally related to genuine symptom improvement or merely constitute an epiphenomenon, in that implicit measures may just fluctuate along with the experienced clinical symptoms. Another important question is whether certain implicit measures can predict symptom onset.

INFERENCEAL CONFUSION: THE DUBIOUS CONSTRUCT VALIDITY OF IMPLICIT MEASURES

The popularity of paradigms to obtain implicit measures, and in particular of the IAT, no doubt derives in large part from the discipline-wide usage of cognitive science methods and focus on indirect measurement procedures for cognition, as in the domains of implicit learning (see Stadler & Frensch, 1998) and implicit memory (see Roediger & Geraci, 2005). But it also rests in part on the promises that are made about what these paradigms can reveal. For example, the website of Project Implicit (<http://projectimplicit.net/index.html>) states that “Project Implicit investigates thoughts and feelings outside of conscious awareness and control,” and the website of Project Implicit Mental Health (<https://implicit.harvard.edu/implicit/user/pimb/background.html>) states that

“implicit measures assess implicit attitudes, identities and beliefs that people are either unwilling or unable to report.” These claims go beyond the definition provided by De Houwer et al. (2009), who have reserved and defined the term “implicit” for the measurement outcome, not for a mental construct (e.g., attitude or belief), and who state that “the implicitness of a measure says little about how the underlying attribute is represented. For instance, it is difficult to determine whether different attributes underlie implicit and explicit measures or whether both measures reflect the same attributes under different conditions” (p. 351). This more narrow definition by De Houwer and colleagues is much more warranted, given the current state of empirical evidence

These statements that appear on the websites about implicit measures are also reflected in research applying the IAT. For example, Asendorpf, Banse, and Mücke (2002, p. 392) concluded, “Just as free associations in psychoanalytic settings provide a window to the unconscious, implicit association tests provide another, probably more reliable window.” Often the measurement outcome of an IAT or APP has been equated with a mental construct, such as an implicit attitude or belief. Examples include “implicit attractiveness beliefs” (e.g., Buhlmann, Teachman, & Kathmann, 2011) and “implicit self-esteem” (e.g., Franck, De Raedt, Dereu, & Van den Abbeele, 2007). Even our own early work made references to attitudes—that is, “implicit attitudes towards high-fat foods” (Roefs & Jansen, 2002). When the measurement outcome of a certain paradigm is equated with a mental construct, every time evidence is obtained with a certain paradigm (e.g., an IAT effect), this is interpreted as evidence for the existence of the mental construct (e.g., an implicit attitude) (De Houwer, 2011). Although such claims are provocative, they are not backed by empirical evidence regarding the implicitness of these measures, let alone regarding the existence of a separate implicit construct (e.g., an “implicit attitude”). Despite implicit measures’ not presupposing introspection, this does not imply that the constructs reflected by implicit measures are indeed unconscious/unavailable for introspection. In other words, though the participant is not asked to engage in introspection when performing an IAT-like measure, this does not mean that the construct the IAT measurement outcome is taken to reflect is not *available* for introspection (Gawronski, LeBel, & Peters, 2007).

VALIDITY ISSUES

The IAT and APP derived initial validity from studies showing that these paradigms can reflect affective associations with stimuli that most people normatively evaluate negatively or positively (e.g., Fazio, 2001; Greenwald et al., 1998). Also, providing stronger support for their validity,

both the APP and IAT reflected associations that were newly created in a classical conditioning procedure while participants showed no contingency awareness (Olson & Fazio, 2001, 2002). However, there are also validity problems with these measures. In the worst case, three “interpretation leaps” are made: (1) that measurement outcomes reflect associations; (2) that associations reflect mental constructs such as attitudes, stereotypes, or beliefs; and (3) that these mental constructs are implicit/unconscious. Doubts surround each of these assumptions.

Measurement Outcomes Reflect Associations

Although there is certainly evidence that the IAT and APP can reflect associations, no measure is process-pure (see the arguments of Jacoby, 1991). In the case of the IAT, other sources of effects have been identified (Fiedler, Messner, & Bluemke, 2006), such as salience asymmetries in the target and attribute dimensions (Rothermund & Wentura, 2004), extrapersonal knowledge (Olson & Fazio, 2004), and cognitive abilities (McFarland & Crouch, 2002; Mierke & Klauer, 2003). We briefly consider each of these.

Salience asymmetries arise because unfamiliar and negative information is typically more salient than is familiar and positive information (see Rothermund & Wentura, 2004, for a review of the evidence). In a series of experiments, Rothermund and Wentura (2004) found that participants were faster to respond when the two salient (e.g., old–nonword) and the two nonsalient (e.g., young–word) categories shared a response key than when the salient and nonsalient categories were mixed. Moreover, valence-free variants of the IAT—that is, IAT variants with attribute categories that have no valence (e.g., word–nonword)—correlated with explicit measures of attitudes. This was of course not an expected result, considering that valence-free IAT variants are not informative regarding attitudes at all. Considering that valence associations and salience asymmetries are typically confounded in applications of the IAT, caution is needed in interpreting these results.

Olson and Fazio (2004) proposed another source of IAT effects: “extrapersonal associations.” These are associations that are available in memory but are not relevant to someone’s personal evaluation of a certain attitude object. For example, a person who is allergic to chocolate may know that people generally are very fond of chocolate. This knowledge may affect the person’s performance on an IAT assessing associations with chocolate. Relevant to this point is the surprising finding that African American participants do not show an in-group bias on the IAT (Nosek, Banaji, & Greenwald, 2002), which may be best explained by their knowledge of the often negative portrayal of African Americans in the mass media. Although Olson and Fazio (2004) agree

that an individual necessarily is part of a culture, and that it is difficult to see how an individual would not be influenced by society (see Banaji, 2001), their main argument is that one's personal attitude can deviate from the cultural norm or the attitude of other people. Knowledge about the cultural norm and other people's attitudes can influence IAT performance. Note, however, that Nosek and Hansen (2008) have argued against culture as a contaminating influence on the IAT, and consider the attempt to separate cultural from personal influences on the IAT to be futile (for an elaborate discussion of this issue, see Olson, Fazio, & Han, 2009).

A third identified source of IAT effects is cognitive skill. Supporting the influence of cognitive skill on the IAT effect, one study found that IATs aimed at assessing prejudice and self-esteem correlated with control IATs with nonsocial and irrelevant dimensions (McFarland & Crouch, 2002). Moreover, an IAT that relied on similarities in the superficial stimulus characteristics of stimulus size and color correlated not only with a second administration of this IAT, but also with an IAT assessing flower–insect preferences (Mierke & Klauer, 2003)—a conceptual dimension. This is suggestive of an underlying cognitive skill. In later research, it appeared that the influence of the cognitive skill confound almost disappeared when the alternative scoring algorithm (*D*-measure) proposed by Greenwald and colleagues (2003) was used (Cai, Siriam, Greenwald, & McFarland, 2004). However, most recent evidence suggests that the *D*-measure does not protect completely against the cognitive skill confound, as cognitive control still correlated with IAT performance (Klauer, Schmitz, Teige-Mocigemba, & Voss, 2010), even when the role of attitude was factored out (Siegel, Dougherty, & Huber, 2012). In addition, by experimentally manipulating cognitive control (e.g., by manipulating awareness of the IAT's goal), it was shown that cognitive control is not merely a correlate of IAT effects, but plays a causal role in it (Siegel et al., 2012). This remaining effect of cognitive skill on the IAT effect may be due to qualitative differences among participants in the type of strategy that is employed to “solve” the sorting task, which cannot be corrected for by a scoring algorithm (Fiedler et al., 2006).

One way to explain most of the alternative sources of IAT effects discussed in the preceding paragraphs is to use the similarity account (De Houwer, Geldof, & De Bruycker, 2005). The central idea of this account is that IAT effects are not restricted to reflect associations in semantic memory (as the IAT was devised to test), but that participants can take advantage of any type of similarity in the stimulus dimensions to perform the task as quickly as possible, including similarities in salience and perceptual characteristics. If associations are broadly defined (see Greenwald, Nosek, Banaji, & Klauer, 2005)—that is, not restricted to associations in a *semantic* network—then the similarity account is not

incompatible with the idea that the IAT measures associations. This broadening does, however, quite radically change the conceptualization of the IAT.

Importantly, as Rothermund and Wentura (2004) pointed out, as soon as participants develop a conscious strategy regarding how best to perform the IAT (i.e., how to take advantage of the observed similarity of presented stimuli), the IAT measurement outcome cannot be considered implicit (see also Fiedler et al., 2006). So if participants can easily recode the task by using only one dimension (e.g., pleasant vs. unpleasant) instead of two (e.g., pleasant vs. unpleasant, peaceful vs. violent) to categorize the stimuli in the compatible phase, and can then apply this strategy consciously, it is certainly reasonable to question whether the measurement outcome is implicit in the sense of reflecting spontaneous associations. It is worth noting in this regard that Greenwald et al. (2005) do not take a position regarding the involvement of controlled processes in the strategy use of participants in a typical IAT. Interesting in this respect is the finding in one study that APP effects relied on so-called “embodied” processes (i.e., activation of the corrugator and zygomatic facial muscles; Foroni & Semin, 2012), whereas IAT effects did not. To the extent that the activation of these facial muscles can be interpreted as automatic affective responses, this pattern is also suggestive of more strategic involvement in the IAT.

Associations Reflect Mental Constructs Such as Attitudes and Beliefs

It is often assumed that implicit measures can be considered the implicit equivalents of explicit dysfunctional beliefs (comparable to attitudes in social psychology); for example, implicit measures of “spider–negative” associations presumably reflect spider-related dysfunctional beliefs (e.g., Huijding & de Jong, 2005). However, can dysfunctional beliefs really be reduced to these simpler associations captured by implicit measures? It may turn out that certain constructs, such as self-esteem, cannot be reduced to simple associations with the self (Buhrmester, Blanton, & Swann, 2011). An illustration would be implicit measures of self-esteem, which are often obtained in research on depressive disorders on the assumption that they reflect dysfunctional depressive beliefs (e.g., De Raedt et al., 2006; Gemar et al., 2001). The assessment of simpler associations is of course not necessarily problematic or of less relevance, but considering them to be the implicit equivalent of explicit dysfunctional beliefs is likely to be problematic.

Another issue concerns the stability of the dysfunctional schemas on the one hand and the stability of the implicit measures on the other hand. Cognitive models in psychopathology typically assume that

dysfunctional schemas are deep-rooted and relatively stable (e.g., Beck & D. A. Clark, 1997; Beck et al., 1979, 2004). Valid measures of these schemas should therefore possess a fair amount of stability over time and situations. However, there is a wealth of evidence indicating the opposite—that implicit measures are highly malleable by the momentary context and short-term manipulations (e.g., Karpinski & Hilton, 2001; Roefs et al., 2006; for reviews, see Blair, 2002; Gawronski et al., 2007). For example, an APP measure of valence associations with food was more sensitive to a short-term manipulation emphasizing health versus taste than to weight status (obese vs. healthy weight) of the participants (Roefs et al., 2006). Often self-report measures prove to be less influenced by these experimental manipulations than implicit measures are (for a review, see Gawronski et al., 2007).

The fact that implicit measures are highly specific—that is, very much influenced by short-term circumstances—may partly account for another frequently mentioned problem, their frequent low test–retest reliability (e.g., Fazio & Olson, 2003; LeBel & Paunonen, 2011), posing problems for their validity and replicability. Test–retest reliability for the IAT is moderately good (median $r = .56$; De Houwer et al., 2009; Nosek, Greenwald, & Banaji, 2007), but the APP findings are less consistent. Some studies on the APP have reported low test–retest reliability (.08–.28; Bosson, Swann, & Pennebaker, 2000), whereas another study, which corrected for measurement error, showed substantially higher test–retest reliability (stability estimate = .68; Cunningham, Preacher, & Banaji, 2001).

Similar problems have been reported regarding internal consistency, which is moderately good for the IAT ($r/\alpha = .7\text{--}.9$), but low for the APP (e.g., Bosson et al., 2000: $\alpha = -.16\text{--}.49$; Olson & Fazio, 2003: $r = .04$), except when a correction for measurement error is included ($\alpha = .64$; Cunningham et al., 2001). The lower-than-ideal reliability estimates also contribute to the problem of low convergent validity over implicit measures (e.g., Bosson et al., 2000; Cunningham et al., 2001), which is notably improved when a latent variable approach is used (Cunningham et al., 2001).

The reliability of implicit measures has been the subject of concern in other research domains (e.g., memory) as well. Indeed, Buchner and Wippich (2000) suggest that the frequently reported dissociation of implicit and explicit memory tests, wherein an effect evident on an explicit test is not seen on an implicit test, is very often due to the unreliability of the implicit test. This unreliability concern leads to the recommendation that because individual response latencies fluctuate over trials (due, for example, to fatigue, practice, and sequence effects), researchers should strive to include as many trials per condition as feasible for

response latency measures. In fact, the improved reliability of the IAT over the APP may be caused by strategic influences on the IAT, as well as associations with categories of stimuli instead of with individual stimuli (Olson & Fazio, 2003). Both of these influences would likely lead to less variability over trials. Naturally, the same reasoning could be applied to self-report instruments, where higher reliability may be inflated because participants often have a tendency to be consistent.

Mental Constructs Are Implicit

As discussed previously, De Houwer et al. (2009) restrict the definition of “implicit” to the measurement outcomes and do not extend it to the mental constructs that are supposedly reflected by these outcomes. In general, it appears doubtful whether implicit measures reflect representations to which participants do not have introspective access (Gawronski et al., 2007). Indeed, even when one limits the discussion to the implicitness of the measurement outcomes, there remains much room for discussion. De Houwer and colleagues (De Houwer, 2006, 2009; De Houwer et al., 2009) provide extensive reviews of the degree to which various implicit measures meet the criteria for implicitness, to which the interested reader is referred. It is important to keep in mind that participants are not necessarily unaware of what is being assessed (e.g., Monteith, Voils, & Ashburn-Nardo, 2001) and are not necessarily unable to control responses (Gawronski et al., 2007), especially when they have experience with, for example, the IAT (e.g., Fiedler & Bluemke, 2005). This effect of IAT experience on controllability poses obvious problems for designs with repeated testing, especially when combined with the observation that the IAT effect generally decreases with repeated administration (e.g., Huijding & de Jong, 2007; Robusto, Cristante, & Vianello, 2008).

PROS AND CONS OF A PARADIGM APPROACH TO RESEARCH

In a paradigm approach, much research is focused on the specifics of a certain paradigm. Currently we seem to have the situation that a behavioral effect (i.e., a difference in response latencies) obtained with a certain indirect measurement procedure (e.g., the IAT or APP) is equated with a mental process or construct (e.g., an unconscious attitude or dysfunctional belief). Although IAT and APP effects are often interpreted as reflecting attitudes to which participants lack introspective access, the empirical evidence reviewed thus far raises doubts about this assertion. The obvious risk, then, is that, in attempting to validate theoretical

models by findings from a specific paradigm, a researcher may not have the right experimental proxy for the mental construct of interest. Put another way, other mental processes/constructs may well also be able to account for the obtained effects (De Houwer, 2011). Yet it is true that this paradigm-oriented research approach dominates most of experimental psychology (Meiser, 2011). Although Meiser admits that it is a typical characteristic of "normal science," he points to the danger of limiting research to the specific features of a certain paradigm without sufficient attention to converging evidence from other paradigms or unifying models that integrate findings from different paradigms.

On the positive side, experimental paradigms have the important function of broadening the range of phenomena that can be measured, and consequently can and do lead to new theories (Greenwald, 2012; Meiser, 2011). A highly illustrative example, provided by Meiser, concerns the four-card Wason selection task (Wason, 1966). Although the Wason task was originally designed to study deductive reasoning, it became evident that the task actually does not assess deductive reasoning. Nevertheless, research with the task continued and inspired new theoretical and empirical approaches. Just one example concerns the effects of content on human reasoning (Cheng & Holyoak, 1985). Individuals use pragmatic reasoning schemas to solve the card selection task. That is, people can solve the task better when it relates to rules in realistic situations with which they have experience.

When we apply this realization to the topic of this chapter, we conclude that although implicit measures may not reflect dysfunctional beliefs or attitudes to which people lack introspective access, this does not mean that research with the IAT and APP is meaningless. As became evident from the review of findings in the field of psychopathology research (see also Roefs et al., 2011), implicit measures have shown incremental predictive validity, and outcomes with these paradigms have sometimes yielded findings that were the opposite of those obtained with explicit measures (e.g., self-esteem questionnaires vs. implicit measures of associations with the self). So these paradigms reflect something potentially interesting and worthy of further study. Put simply, such dissociations are provocative and lead us to new interpretations and to new research approaches.

WHERE TO NEXT?

To avoid the problems created by treating implicit measures as the implicit equivalents of explicit attitudes and beliefs, researchers should in their interpretation of results remain true to what has actually been

assessed by the experimental task. In a sense, it is quite surprising that researchers did not initially do that, considering that they have done so more often in research that adapted paradigms from cognitive psychology directed at other cognitive processes and constructs, such as attention (e.g., Field, Munafo, & Franken, 2006) and memory (e.g., Mitte, 2008). In research on attention bias, for instance, effects in paradigms such as the dot-probe task (e.g., Bradley, Mogg, & Lee, 1997) have been interpreted in terms of attention processes, and not directly extended to attitudes or beliefs. Thus the interpretation is closer to the phenomena that are assessed by the paradigm, involving less inference to mental constructs.

Rather than making leaps from measurement outcomes, via associations, to implicit dysfunctional beliefs or attitudes, it would seem better to restrict interpretations of APP and IAT results to a tendency to associate certain concepts more strongly than others, which might be termed "association bias." It remains for further research to ascertain which paradigm is best suited to capturing such an association bias. It may be that the APP comes closer to assessing spontaneous associations than the IAT, as the IAT is an inherently more complex task that elicits strategy use in at least some participants. So, from a cognitive perspective, on the one hand researchers should focus on the paradigm that is best capable of assessing the theoretical construct of interest, and on the other hand they should strive for convergence of results across various paradigms. It should be noted that convergence may be expected only when paradigms are structurally similar, not just similar in the goal of obtaining implicit measures. When we consider De Houwer's (2003) structural analysis of the APP and the IAT (among other paradigms), it is no longer surprising that the results obtained in these paradigms do not necessarily correlate highly.

To be able to design the best paradigm to study dysfunctions in cognitions, cognitive models should be formulated in specific terms. As we have noted at the outset, much of the research with implicit measures has relied on Beck's schema model (e.g., Beck & D. A. Clark, 1997; Beck et al., 1979, 2004). As Van den Hout (1999) has argued, the heuristic value of the belief concept is undeniable, and it is undoubtedly true that dysfunctional beliefs play an important role in psychopathology. Manipulations of these beliefs lead to changes in psychopathology. To gain a more profound understanding of specific cognitive dysfunctions in psychopathology, though, it certainly will be beneficial to specify cognitive models of psychopathology in greater detail, beyond just proposing a dysfunctional schema. As Segal (1988) has argued, it will be important to learn from theoretical progress regarding knowledge representations in cognitive science, and to incorporate this learning in the

conceptualization of dysfunctional schemas. While it is true that such models do already exist for certain disorders, as noted earlier in this chapter, they have not formed the basis for research with implicit measures. When an explanation is more specific about the precise mental process that is disturbed, more focused paradigms can be developed that also permit more effective use of converging measures. Although this issue is certainly not unique to the domain of psychopathology, it is also certainly very relevant in this domain. Good paradigms tend to develop, and to produce reliable data, when they are based on well-specified theories (Strauss & Smith, 2009).

One way to design more specific theoretical models is to combine the cognitive approach with a functional approach, as recently argued by De Houwer (2011):

A functional approach to psychology is the practice of defining behavioral effects exclusively in terms of elements in the environment [p. 204]. . . . The functional approach is concerned mainly with the question of when certain elements in the environment influence behavior. The cognitive approach addresses the question of how those elements in the environment influence behavior. It does so by describing the mental processes/representations that are assumed to mediate the effect of the environment on behavior [p. 205]. . . . The more we know about when a behavioral effect occurs, the more precise we can be about the mental constructs that mediate this effect [p. 205]. . . . Vice versa, mental explanations can generate new hypotheses about the conditions under which behavioral effects occur [pp. 205–206].

For example, if it is consistently found (in different paradigms) that participants with depressive disorder when required to give a speeded response, respond faster when the concepts of “me” and “failure” are paired than when the concepts of “me” and “success” are paired (functional description), this may provide boundary conditions for the structure of the mental representation of this knowledge (cognitive description).

As has become evident in the part of this chapter on validity issues, it seems safe to conclude that personal attitudes or beliefs can be a source of IAT and APP effects. In contrast, the reverse inference is considerably more problematic: IAT and APP effects cannot with certainty be interpreted as evidence for the mental constructs of (unconscious) attitudes or beliefs, or for the existence of latent schemas. Yet it is certainly the case that evidence obtained with these paradigms has been interpreted as evidence for theories postulating unconscious attitudes, which is of course problematic if effects obtained with these paradigms do not prove to be valid mental proxies. We believe that future research

would benefit from a combined functional and cognitive approach. This approach would start by describing and studying IAT and APP effects in functional terms, which could then lead to a more specific description of the mental constructs—the cognitive processes—potentially responsible for these effects.

Table 13.1 (inspired by De Houwer's [2011] Table 1) makes it clear that the cognitive and functional approaches yield quite different descriptions. The same approach could be taken for other experimental tasks that are often used to test cognitive models of psychopathology, probing functions like attention and memory. When this approach is taken, diverging results from different paradigms—constituting a lack of convergent validity in a cognitive approach—are not problematic, and are even to be expected, given that tasks are never process-pure. What results is a more detailed and complete *functional picture* of a disorder. Convergence can then be striven for at this functional level. This functional perspective can then be used to form the basis of a well-informed specific cognitive model for different forms of psychopathology.

TABLE 13.1. Cognitive versus Functional Approaches to IAT and APP

Element	Cognitive approach	IAT functional approach	APP functional approach
Definition	Differential response speed due to associations in memory/unconscious attitudes or dysfunctional beliefs	Slower response speed when less similar categories are combined under one response button	Slower response speed with less similar prime-target combinations
Explanandum ^a	IAT and APP effect	Differential response speed in the two combined phases	Differential response speed with different types of prime-target combinations
Explanans ^b	Unconscious attitudes and dysfunctional beliefs	Combinations of categories	Prime-target combinations

^aExplanandum, what needs to be explained.

^bExplanans, that which is used to explain.

CONCLUSION

In this chapter, we have proposed that researchers should refrain from straying too far from the phenomena when making interpretations of what is actually measured in the experimental task. Researchers using paradigms like the IAT and the APP tend to go from measurement outcome via association to implicit attitude or dysfunctional belief, assuming that this is a shorter inferential route than is actually the case. It is important to note that this has not been the case so much with paradigms capturing other constructs that are deemed of relevance in cognitive theories of psychopathology, such as attention and memory biases. Restricting the interpretation of IAT and APP results to an *association bias* is probably more justified. Taking this to the next level is the functional approach (De Houwer, 2011), thereby preventing problems inherent with a cognitive approach that uses behavioral proxies for mental constructs. The cognitive approach clearly will benefit from the detailed specific results from the functional approach, and the functional approach will be inspired by cognitive theories.

KEY POINTS

FOR PRACTITIONERS

- Implicit measures are useful as a research tool for theory development, but are not ready for individual diagnosis or clinical practice, due to incomplete knowledge regarding their validity and reliability and to their frequent low reliability.
- However, implicit measures may *eventually* provide useful insights into clinical disorders that may call for changes in therapy.

FOR RESEARCHERS

- Researchers should in their interpretation of results remain true to what has actually been assessed by the experimental task. Therefore, the interpretation of implicit measures should be restricted to *association bias*, instead of extending it to implicit attitudes or beliefs.
- There is insufficient empirical evidence regarding the existence of a separate implicit construct (e.g., an "implicit attitude"), nor is there enough evidence regarding the implicitness of these measures.
- A behavioral effect (e.g., an IAT effect) should not be equated with a mental process or construct (e.g., an unconscious attitude or dysfunctional belief).
- Psychopathology research that departs from a cognitive perspective is

likely to benefit from keeping an eye on the state-of-the-art theorizing and methodologies from cognitive science.

FOR STUDENTS AND EDUCATORS

- Avoid gaining/teaching an oversimplified view of implicit measures, but instead include discussion on the validity of implicit measures.
- Think carefully, or have students think carefully, about what implicit measures really reflect. Do not simply accept the interpretations that are provided in research articles.

REFERENCES

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, *83*, 380–393.
- Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger & J. S. Nairne (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 117–150). Washington, DC: American Psychological Association.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.
- Beck, A. T., & Clark, D. A. (1997). An information processing model of anxiety: Automatic and strategic processes. *Behaviour Research and Therapy*, *35*, 49–58.
- Beck, A. T., Freeman, A., & Davis, D. D. (Eds.). (2004). *Cognitive therapy of personality disorders* (2nd ed.). New York: Guilford Press.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*, 242–261.
- Boring, E. G. (1953). A history of introspection. *Psychological Bulletin*, *50*, 169–189.
- Bosson, J. K., Swann, W. B., & Penebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology*, *79*, 631–643.
- Bradley, B. P., Mogg, K., & Lee, S. C. (1997). Attentional biases for negative information in induced and naturally occurring dysphoria. *Behaviour Research and Therapy*, *35*, 911–927.
- Buchner, A., & Wippich, W. (2000). On the reliability of implicit and explicit memory measures. *Cognitive Psychology*, *40*, 227–259.
- Buhlmann, U., Teachman, B. A., & Kathmann, N. (2011). Evaluating implicit attractiveness beliefs in body dysmorphic disorder using the go/no-go

- association task. *Journal of Behavior Therapy and Experimental Psychiatry*, 42, 192–197.
- Buhrmester, M. D., Blanton, H., & Swann, W. B. (2011). Implicit self-esteem: Nature, measurement, and a new way forward. *Journal of Personality and Social Psychology*, 100(2), 365–385.
- Cai, H., Siriam, N., Greenwald, A. G., & McFarland, S. G. (2004). The Implicit Association Test's D measure can minimize a cognitive skill confound: Comment on McFarland and Crouch (2002). *Social Cognition*, 22, 673–684.
- Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391–416.
- Clark, D. A., Beck, A. T., & Alford, B. A. (1999). *Scientific foundations of cognitive theory and therapy of depression*. New York: Wiley.
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 69–93). New York: Guilford Press.
- Clerkin, E. M., & Teachman, B. A. (2009). Automatic and strategic measures as predictors of mirror gazing among individuals with body dysmorphic disorder symptoms. *Journal of Nervous and Mental Disease*, 197, 589–598.
- Craeynest, M., Crombez, G., Deforche, B., Tanghe, A., & De Bourdeaudhuij, I. (2008). The role of implicit attitudes towards food and physical activity in the treatment of youth obesity. *Eating Behaviors*, 9, 41–51.
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measurement: Consistency, stability, and convergent validity. *Psychological Science*, 12, 163–170.
- De Houwer, J. (2003). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219–244). Mahwah, NJ: Erlbaum.
- De Houwer, J. (2006). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction* (pp. 11–28). Thousand Oaks, CA: Sage.
- De Houwer, J. (2009). Comparing measures of attitudes at the functional and procedural level: Analysis and implications. In R. E. Petty, R. H. Fazio, & P. Brinol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 361–390). New York: Erlbaum.
- De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science*, 6, 202–209.
- De Houwer, J., Geldof, T., & De Bruycker, E. (2005). The Implicit Association Test as a general measure of similarity. *Canadian Journal of Experimental Psychology*, 59, 228–239.
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, 135, 347–368.
- De Raedt, R., Schacht, R., Franck, E., & De Houwer, J. (2006). Self-esteem and

- depression revisited: Implicit positive self-esteem in depressed patients? *Behaviour Research and Therapy*, 44, 1017–1028.
- Fazio, R. H. (2001). On the automatic evaluation of associated evaluations: An overview. *Cognition and Emotion*, 15, 115–141.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, 54, 297–327.
- 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.
- Fazio, R. H., & Towles-Schwen, T. (1999). The MODE model of attitude-behavior processes. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 97–116). New York: Guilford Press.
- 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.
- Field, M., Munafo, M. R., & Franken, I. H. (2009). A meta-analytic investigation of the relationship between attentional bias and subjective craving in substance abuse. *Psychological Bulletin*, 135, 589–607.
- Froni, F., & Semin, G. R. (2012). Not all implicit measures of attitudes are created equal: Evidence from an embodiment perspective. *Journal of Experimental Social Psychology*, 48, 424–427.
- Franck, E., De Raedt, R., Dereu, M., & Van den Abbeele, D. (2007). Implicit and explicit self-esteem in currently depressed individuals with and without suicidal ideation. *Journal of Behavior Therapy and Experimental Psychiatry*, 38, 75–85.
- Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us?: Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science*, 2, 181–193.
- Gemar, M. C., Segal, Z. V., Sagrati, S., & Kennedy, S. J. (2001). Mood-induced changes on the Implicit Association Test in recovered depressed patients. *Journal of Abnormal Psychology*, 110, 282–289.
- Greenwald, A. G. (2012). There is nothing so theoretical as a good method. *Perspectives on Psychological Science*, 7, 99–108.
- 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., 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.
- Greenwald, A. G., Nosek, B. A., Banaji, M. R., & Klauer, K. C. (2005). Validity of the salience asymmetry interpretation of the Implicit Association Test:

- Comment on Rothermund and Wentura (2004). *Journal of Experimental Psychology: General*, 134, 420–425.
- Grumm, M., Erbe, K., von Collani, G., & Nestler, S. (2008). Automatic processing of pain: The change of implicit pain associations after psychotherapy. *Behaviour Research and Therapy*, 46, 701–714.
- Haefffel, G. J., Abramson, L. Y., Brazy, P. C., Shah, J. Y., Teachman, B. A., & Nosek, B. A. (2007). Explicit and implicit cognition: A preliminary test of a dual-process theory of cognitive vulnerability to depression. *Behaviour Research and Therapy*, 45, 1155–1167.
- Hofmann, W., Friese, M., & Roefs, A. (2009). Three ways to resist temptation: The independent contributions of executive attention, inhibitory control, and affect regulation to the impulse control of eating behavior. *Journal of Experimental Social Psychology*, 45, 431–435.
- Hofmann, W., Rauch, W., & Gawronski, B. (2007). And deplete us not into temptation: Automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behavior. *Journal of Experimental Social Psychology*, 43, 497–504.
- Houben, K., & Wiers, R. W. (2006). Assessing implicit alcohol associations with the Implicit Association Test: Fact or artifact? *Addictive Behaviors*, 31, 1346–1362.
- Houben, K., & Wiers, R. W. (2008). Implicitly positive about alcohol?: Implicit positive associations predict drinking behavior. *Addictive Behaviors*, 33, 979–986.
- Huijding, J., & de Jong, P. J. (2005). A pictorial version of the extrinsic affective Simon task: Sensitivity to generally affective and phobia-relevant stimuli in high and low spider fearful individuals. *Experimental Psychology*, 52, 289–295.
- Huijding, J., & de Jong, P. J. (2006). Specific predictive power of automatic spider-related affective associations for controllable and uncontrollable fear responses toward spiders. *Behaviour Research and Therapy*, 44, 161–176.
- Huijding, J., & de Jong, P. J. (2007). Beyond fear and disgust: The role of (automatic) contamination-related associations in spider phobia. *Journal of Behaviour Therapy and Experimental Psychiatry*, 38, 200–211.
- Ingram, R. E. (1984). Toward an information-processing analysis of depression. *Cognitive Therapy and Research*, 8, 443–477.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30, 513–541.
- Joormann, J., Yoon, K. L., & Zetsche, U. (2007). Cognitive inhibition in depression. *Applied and Preventive Psychology*, 12, 128–139.
- Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology*, 81, 774–788.
- Klauer, K. C., Schmitz, F., Teige-Mocigemba, S., & Voss, A. (2010). Understanding the role of executive control in the Implicit Association Test: Why flexible people have small IAT effects. *Quarterly Journal of Experimental Psychology*, 63, 595–619.
- LeBel, E. P., & Paunonen, S. V. (2011). Sexy but often unreliable: The impact

- of unreliability on the replicability of experimental findings with implicit measures. *Personality and Social Psychology Bulletin*, 37, 570–583.
- MacLeod, C. (1993). Cognition in clinical psychology: Measures, methods or models? *Behaviour Change*, 10, 169–195.
- McFarland, S. G., & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition*, 20, 483–510.
- McNally, R. J. (2001). On the scientific status of cognitive appraisal models of anxiety disorder. *Behaviour Research and Therapy*, 39, 513–521.
- Meiser, T. (2011). Much pain, little gain?: Paradigm-specific models and methods in experimental psychology. *Perspectives on Psychological Science*, 6, 183–191.
- Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology*, 85, 1180–1192.
- Mitte, K. (2008). Memory bias for threatening information in anxiety and anxiety disorders: A meta-analytic review. *Psychological Bulletin*, 134, 886–911.
- Monteith, M. J., Voils, C. I., & Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial bias. *Social Cognition*, 19, 395–417.
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, 132, 297–326.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration website. *Group Dynamics*, 6, 101–115.
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behaviour* (pp. 265–292). New York: Psychology Press.
- Nosek, B. A., & Hansen, J. J. (2008). The associations in our heads belong to us: Searching for attitudes and knowledge in implicit evaluation. *Cognition and Emotion*, 22, 553–594.
- Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, 12, 413–417.
- Olson, M. A., & Fazio, R. H. (2002). Implicit acquisition and manifestation of classically conditioned attitudes. *Social Cognition*, 20, 89–104.
- Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science*, 14, 636–639.
- Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extrapersonal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology*, 86, 653–667.
- Olson, M. A., Fazio, R. H., & Han, H. A. (2009). Conceptualizing personal and extrapersonal associations. *Social and Personality Psychology Compass*, 3, 152–170.
- Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology*, 44, 29–45.

- Robusto, E., Cristante, F., & Vianello, M. (2008). Assessing the impact of replication on Implicit Association Test effect by means of the extended logistic model for the assessment of change. *Behavior Research Methods, 40*, 954–960.
- Roediger, H. L., III, & Geraci, L. (2005). Implicit memory tasks in cognitive research. In A. Wenzel & D. Rubin (Eds.), *Cognitive methods and their application to clinical research* (pp. 129–151). Washington, DC: American Psychological Association.
- Roefs, A., Huijding, J., Smulders, F. T. Y., MacLeod, C. M., de Jong, P. J., Wiers, R. W., et al. (2011). Implicit measures of association in psychopathology research. *Psychological Bulletin, 137*, 149–193.
- Roefs, A., & Jansen, A. (2002). Implicit and explicit attitudes toward high-fat foods in obesity. *Journal of Abnormal Psychology, 111*, 517–521.
- Roefs, A., Quaedackers, L., Werrij, M. Q., Wolters, G., Havermans, R., Nederkoorn, C., et al. (2006). The environment influences whether high-fat foods are associated with palatable or with unhealthy. *Behaviour Research and Therapy, 44*, 715–736.
- Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test: Dissociating salience from associations. *Journal of Experimental Psychology: General, 133*, 139–165.
- Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Erlbaum.
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54*, 93–105.
- Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *American Journal of Evaluation, 22*, 127–160.
- Segal, Z. V. (1988). Appraisal of the self-schema construct in cognitive models of depression. *Psychological Bulletin, 103*, 147–162.
- Seibt, B., Häfner, M., & Deutsch, R. (2007). Prepared to eat: How immediate affective and motivational responses to food cues are influenced by food deprivation. *European Journal of Social Psychology, 37*, 359–379.
- Siegel, E. F., Dougherty, M. R., & Huber, D. E. (2012). Manipulating the role of cognitive control while taking the Implicit Association Test. *Journal of Experimental Social Psychology, 48*, 1057–1068.
- Stadler, M. A., & Frensch, P. A. (Eds.). (1998). *Handbook of implicit learning*. Thousand Oaks, CA: Sage.
- Strauss, M. E., & Smith, G. T. (2009). Construct validity: Advances in theory and methodology. *Annual Review of Clinical Psychology, 5*, 1–25.
- Teachman, B. A., Marker, C. D., & Smith-Janik, S. B. (2008). Automatic associations and panic disorder: Trajectories of change over the course of treatment. *Journal of Consulting and Clinical Psychology, 76*, 988–1002.
- Teachman, B. A., Smith-Janik, S. B., & Saporito, J. (2007). Information processing biases and panic disorder: Relationships among cognitive and symptom measures. *Behaviour Research and Therapy, 45*, 1791–1811.

- Teachman, B. A., & Woody, S. R. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. *Journal of Abnormal Psychology, 112*, 100–109.
- Thush, C., Wiers, R. W., Ames, S. L., Grenard, J. L., Sussman, S., & Stacy, A. W. (2007). Apples and oranges?: Comparing indirect measures of alcohol-related cognition predicting alcohol use in at-risk adolescents. *Psychology of Addictive Behaviors, 21*, 587–591.
- Van den Hout, M. (1999). Armies of idiots and idiosyncrasies: On reductions in experimental psychopathology. *Behaviour Research and Therapy, 37*, S135–S145.
- Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology* (pp. 135–151). Harmondsworth, UK: Penguin Books.
- Wiers, R. W., van de Luitgaarden, J., van den Wildenberg, E., & Smulders, F. T. (2005). Challenging implicit and explicit alcohol-related cognitions in young heavy drinkers. *Addiction, 100*, 806–819.
- Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120*, 3–24.
- Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. (1997). *Cognitive psychology and emotional disorders*. Chichester, UK: Wiley.