In K. C. Klauer, A. Voss, & C. Stahl (Eds.), Cognitive methods in social psychology. New York: Guilford Press.

# 3

# Response Interference Tasks as Indirect Measures of Automatic Associations

BERTRAM GAWRONSKI
ROLAND DEUTSCH
RAINER BANSE

A large proportion of research in social psychology uses self-report measures to assess socially relevant constructs, such as attitudes, stereotypes, self-concepts, and self-esteem. Over the past century, self-report measures have provided important insights that built the foundation for many prominent theories in the field. At the same time, there have been persistent concerns that self-report measures are suboptimal research tools for at least two reasons. First, responses on self-report measures are susceptible to self-presentation, which may distort measurement outcomes in socially sensitive domains (Crowne & Marlowe, 1960; Paulhus, 1984). Second, there is consensus that many psychological processes operate outside of conscious awareness, which undermines the suitability of self-report measures to assess mental contents that are introspectively inaccessible (Nisbett & Wilson, 1977). To overcome these problems, psychologists have developed various indirect measurement procedures that (1) reduce participants' ability to control their responses, and (2) do not require introspection for the assessment of mental contents.

This chapter reviews a particularly influential class of indirect measurement procedures, namely, experimental paradigms based on response interference (RI).[1] For this purpose, we first provide a general context for our review of RI tasks by defining a number of relevant concepts and briefly describing the history of RI tasks in social psychology. After explaining the basic logic of RI, we then review different kinds of measures based on RI, including general information about their procedures, implementation, and scoring. We conclude our chapter with a discussion of general issues regarding the interpretation of RI effects and a pragmatically oriented comparison of the reviewed tasks.





## Basic Concepts and Terminology

The tasks reviewed in this chapter—including the sequential priming tasks reviewed by Spruyt et al. (Chapter 2, this volume)—are often referred to as *implicit measures*. A common source of confusion with regard to this label is the meaning of the term *implicit*. Whereas some researchers use the term *implicit* to describe a particular characteristic of the measurement procedures (e.g., Fazio & Olson, 2003), others use it to describe the constructs assessed by a particular class of measures (e.g., Greenwald & Banaji, 1995). The first meaning is based on the fact that RI and sequential priming tasks provide indicators of psychological attributes (e.g., attitudes) without having to ask participants to report the desired information verbally (e.g., Fazio & Olson, 2003). In contrast, the second meaning is based on the fact that RI and sequential priming tasks do not require introspection and may therefore reflect psychological attributes that are unconscious or introspectively inaccessible (e.g., Banaji, 2001).

In this chapter, we follow suggestions by De Houwer (2006) to use the terms *direct* and *indirect* to describe features of measurement procedures, and the terms *implicit* and *explicit* to describe features of the psychological attributes assessed by different measurement procedures. In terms of De Houwer's distinction, RI tasks fall under the category of indirect measurement procedures, in that participants are not asked directly to provide a self-assessment of the relevant psychological attribute (e.g., attitude, self-esteem). De Houwer further suggests use of the term *implicit* to describe constructs that influence task performance in an automatic fashion. In terms of Bargh's (1994) "four horsemen of automaticity," this influence may occur unconsciously, unintentionally, efficiently, and/or uncontrollably (see also Moors & De Houwer, 2006).

Another important issue concerns terminology for the particular con-

[1]Another influential type of indirect measurement procedures is sequential priming tasks (e.g., Bargh, Chaiken, Raymond, & Hymes, 1996; Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Payne, Cheng, Govorun, & Stewart, 2005; Wittenbrink, Judd, & Park, 1997). These tasks are discussed in more detail in a separate chapter (see Spruyt, Gast, & Moor, Chapter 2, this volume).

structs assessed by RI tasks. A common conceptualization refers to these constructs as *associations*. This convention is based on recent definitions of major social-psychological constructs as associations between two concepts (e.g., Greenwald et al., 2002). For instance, the term *attitude* has been defined as the association between an object and its evaluation in memory (Fazio, 2007). Along the same lines, *self-esteem* can be defined as an association between the self and its evaluation. With regard to nonevaluative constructs, *stereotypes* can be conceptualized as associations between a social category and a stereotypically related attribute (e.g., women–warm), just as *self-concept* can be defined as the association between the self and its attributes (e.g., self–extraverted). Based on these conceptualizations, we describe the constructs assessed by RI tasks as *automatic associations*. Note, however, that the presumed automatic nature of these associations is not guaranteed by the indirect nature of the measurement procedures. Whether the associations assessed by RI tasks are indeed unconscious, efficient, unintentional, and/or uncontrollable is an empirical question that has to be tested by means of carefully designed studies (see De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009; Gawronski, Hofmann, & Wilbur, 2006). Thus, our use of the term *automatic association* simply reflects the general purpose of RI tasks to assess associations that are in fact automatic rather than a methodological necessity that is logically implied by the indirect nature of RI tasks. We return to this issue in the concluding sections of this chapter when we discuss general issues regarding the interpretation of measurement scores.

## A Brief History

The seminal work that has set the stage for the use of RI tasks in social psychology is Fazio, Sanbonmatsu, Powell, and Kardes's (1986) adaptation of sequential priming to study the automatic activation of attitudes, defined as the association between objects and their summary evaluations in memory (Fazio, 2007). This line of research led to a number of groundbreaking findings in the attitudes literature (for a review, see Olson & Fazio, 2009), and Fazio et al.'s (1986) evaluative priming task still represents one of the most frequently used indirect measures in social psychology. About a decade later, the popularity of indirect measures surged when Greenwald, McGhee, and Schwartz (1998) presented their widely known Implicit Association Test (IAT). Different from earlier research on automatic attitude activation, the development of the IAT was inspired by the notion of implicit memory, in that the IAT has been assumed to assess unconscious associations that are inaccessible to introspection (Banaji, 2001). Subsequent theoretical work emphasized the potential of indirect measures in assessing highly stable associations that have their roots in early socialization experiences. The general assumption in these theories is that indirect measures assess old



associations that have not been erased from memory, even when newly formed associations lead to observable changes in verbal self-reports (e.g., Petty, Tormala, Briñol, & Jarvis, 2006; Rudman, 2004; Wilson, Lindsey, & Schooler, 2000). More recent theorizing introduced an interpretation that has a stronger resemblance to the aforementioned work on automatic attitude activation, arguing that performance on indirect measures reflects the momentary accessibility of mental associations that may have chronic and situational sources (e.g., Gawronski & Bodenhausen, 2006; Strack & Deutsch, 2004). This interpretation expands on earlier work using similar methods to assess the accessibility of single concepts (e.g., lexical decision tasks; see Förster & Liberman, 2007); however, the newly developed indirect measures go beyond these earlier methods by targeting the accessibility of associations between two (or more) concepts.[2]

Even though there is still controversy about the psychological nature of the associations assessed by RI tasks (e.g., Arkes & Tetlock, 2004; De Houwer, 2006; Gawronski, 2009), their usefulness has been demonstrated in numerous studies showing that they predict judgments and behavior over and above standard self-report measures (for reviews, see Friese, Hofmann, & Schmitt, 2008; Perugini, Richetin, & Zogmaister, 2010). Taken together, these studies have shown that RI tasks tend to outperform self-report measures in the prediction of overt behavior for particular types of behavior (e.g., spontaneous behavior; Asendorpf, Banse, & Mücke, 2002), under particular circumstances (e.g., when cognitive resources are depleted; Hofmann, Rauch, & Gawronski, 2007), and for individuals with particular processing styles (e.g., with a preference for intuitive decisions; Richetin, Perugini, Adjali, & Hurling, 2007). In addition, the use of RI tasks has provided significant insights into the processes underlying the formation and change of mental associations (for a review, see Gawronski & Sritharan, 2010), which has inspired a number of theoretical advances in this area (e.g., Gawronski & Bodenhausen, 2006; Petty, Briñol, & DeMarree, 2007).

## RI Tasks: Logic and Procedures

As outlined earlier, RI tasks represent one of the most influential types of indirect measurement procedures. The basic logic of RI tasks is that an ostensibly irrelevant or to-be-ignored stimulus elicits an impulsive or prepotent response tendency that may either facilitate or inhibit quick and/or accurate responses in the primary task. A useful example to illustrate the logic of RI is the Stroop color-naming task (Stroop, 1935), in which participants are asked to name the ink color of a word as quickly as pos-

[2]As outlined later in this chapter, some indirect measures (e.g., IAT) assess associations among multiple concepts.

sible. The critical items in this task are words that themselves are color names. On these items, participants usually show better performance when the ink color of the word corresponds to the meaning of the word (e.g., the word *red* written in red ink) than when ink color and word meaning do not correspond to one another (e.g., the word *red* written in blue ink). These differences in performance can be explained by the influence of two independent response tendencies elicited by the ink color and the semantic meaning of the stimulus. For instance, the word *red* written in blue ink may elicit two response tendencies that interfere with a quick and accurate response to that stimulus, namely, a response tendency to say "red," elicited by the word meaning, and a response tendency to say "blue," elicited by the ink color. Conversely, the word *red* written in red ink may elicit two identical response tendencies that facilitate quick and accurate responses. Put differently, the first case results in two response tendencies that have antagonistic effects, whereas the latter case results in two response tendencies that have synergistic effects. Performance (e.g., speed, accuracy) tends to be facilitated on trials with synergistic response tendencies, but impaired on trials with antagonistic response tendencies.

RI processes constitute the basis for the large majority of experimental paradigms to assess automatic associations (Gawronski, Deutsch, LeBel, & Peters, 2008). These measures include the IAT (Greenwald et al., 1998) and its derivates (e.g., Karpinski & Steiman, 2006; Olson & Fazio, 2004; Penke, Eichstaedt, & Asendorpf, 2006; Rothermund, Teige-Mocigemba, Gast, & Wentura, 2009; Sriram & Greenwald, 2009; Teige-Mocigemba, Klauer, & Rothermund, 2008), approach–avoidance tasks (e.g., Brendl, Markman, & Messner, 2005; Chen & Bargh, 1999; Schnabel, Banse, & Asendorpf, 2006), the Go/No-Go Association Task (GNAT; Nosek & Banaji, 2001), the Extrinsic Affective Simon Task (EAST; De Houwer, 2003), and the Action Interference Paradigm (Banse, Gawronski, Rebetez, Gutt, & Morton, 2010).[3] In the following sections, we provide a method-oriented overview of these tasks, including information about their task structure, flexibility, procedural details, and scoring.

## Implicit Association Test

### *Task Structure*

The critical trials of the IAT include two binary categorization tasks, which are combined in either an association-congruent or an association-incongruent manner. For example, in an IAT to assess evaluative associations regarding black versus white people (e.g., Rudman, Greenwald, Mellott, & Schwartz, 1999), participants are successively presented with positive and negative words and pictures showing black and white faces that have to be classified as positive and negative or as black and white, respectively. In one of the two critical blocks, the two categorization tasks are combined in such a manner that participants have to respond to positive words and pictures of white faces with one key, and to negative words and pictures of black faces with another key. In the other critical block, the two categorization tasks are combined in the reverse manner; that is, participants have to respond to positive words and pictures of black faces with one key, and to negative words and pictures of white faces with another key. The basic idea underlying the IAT is that quick and accurate responses in the task are facilitated when the key mapping in the task is congruent with participants' automatic associations (e.g., black–negative; white–positive) but inhibited when the key mapping in the task is association-incongruent (e.g., white–negative; black–positive). Based on this consideration, the mean difference in participants' performance in the two blocks is typically interpreted as an index of automatic preferences, for instance, preferences for white over black people (or the other way round, depending on the calculation of the difference score).

[3]In addition to these tasks, there is a particular class of sequential priming tasks that is based on RI processes: category priming tasks (e.g., Banaji & Hardin, 1996; Fazio et al., 1986). As sequential priming tasks are extensively reviewed in Chapter 2, we refrain from discussing category priming tasks in this chapter and refer interested readers to Spruyt et al. (Chapter 2, this volume).



A typical IAT includes a total of five blocks, which are further divided into seven stages (see Table 3.1). Using the aforementioned example, an IAT assessing preferences for white over black people would begin with the first practice block in which participants are asked to categorize pictures of black and white faces as fast and accurately as possible as black versus

**TABLE 3.1. Task Structure of an Implicit Association Test (Greenwald et al., 1998) Designed to Assess Preferences for Whites over Blacks (Race-IAT)**

| | | | Key assignment | | | |
|---|---|---|---|---|---|---|
| | | | Congruent/incongruent block order | | Incongruent/congruent block order | |
| Block | Stage | Trial No. | Left key | Right key | Left key | Right key |
| 1 | 1 | 20 | Negative | Positive | Negative | positive |
| 2 | 2 | 20 | Black | White | White | Black |
| 3 | 3 | 20 | Negative/ black | Positive/ white | Negative/ white | Positive/ black |
| | 4 | 40 | Negative/ black | Positive/ white | Negative/ white | Positive/ black |
| 4 | 5 | 40 | White | Black | Black | White |
| 5 | 6 | 20 | Negative/ white | Positive/ black | Negative/ black | Positive/ white |
| | 7 | 40 | Negative/ white | Positive/ black | Negative/ black | Positive/ white |

white (*initial target–concept discrimination*). In the second practice block, participants are presented with positive and negative words that have to be categorized as pleasant versus unpleasant, again, as quickly and accurately as possible (*associated attribute discrimination*). In the third block, the two categorization tasks are combined, such that participants are presented with words and pictures in alternating order, which have to be categorized with the same key assignments as in the first two blocks (*initial combined task*). For instance, participants may be asked to press a right-hand key every time they see a positive word or a picture of a white person and a left-hand key every time they see a negative word or a picture of a black person. As with the first two blocks, participants are asked to respond as quickly and accurately as possible. The fourth block is almost equivalent to the first block, the only difference being that the key assignment for the two categories is now reversed (*reversed target–concept discrimination*). Finally, the fifth block again combines the two categorization tasks, this time using the key assignments of the second and fourth blocks (*reversed combined task*). In the current example, this would imply that participants have to press a right-hand key every time they see a positive word or a picture of a black person and a left-hand key every time they see a negative word or a picture of a white person.

### Flexibility

The IAT is very flexible and can be applied to assess almost any type of association between pairs of concepts. For instance, by using evaluative attribute dimensions (i.e., pleasant vs. unpleasant) the IAT can be used to assess relative preferences between pairs of objects or categories. Alternatively, the evaluative attribute dimension may be replaced with a semantic dimension to assess semantic associations (e.g., extraverted vs. introverted). The same flexibility applies to the use of target categories, which may include any pair of objects or categories that form meaningful opposites (e.g., male vs. female; Coke vs. Pepsi). Examples of previous IAT applications include prejudice (e.g., Rudman et al., 1999), stereotypes (e.g., Gawronski, Ehrenberg, Banse, Zukova, & Klauer, 2003), attitudes toward consumer products (e.g., Maison, Greenwald, & Bruin, 2004), associations to phobic stimuli (e.g., Teachman & Woody, 2003), the self-concept of personality (e.g., Asendorpf et al., 2002), self-esteem (e.g., Greenwald & Farnham, 2000), motivational associations regarding health-related objects (e.g., Wiers, Van Woerden, Smulders, & De Jong, 2002), and many other kinds of associations (for an overview, see Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005).

The enormous flexibility of the IAT notwithstanding, it is important to note that its basic structure involves pairs of target concepts and pairs of attribute concepts. Thus, IAT effects may be driven by four different kinds of associations (Blanton, Jaccard, Gonzales, & Christie, 2006). For instance, scores in the aforementioned IAT designed to assess preferences for white over black people can be influenced by (1) positive associations to whites, (2) positive associations to blacks, (3) negative associations to whites, and (4) negative associations to blacks. This issue has inspired the development of a number of IAT derivates and other tasks to assess associations to a single concept (see below). Nevertheless, a large body of evidence supports the validity of the IAT, including an impressive number of studies demonstrating the predictive validity of the task (for a meta-analysis, see Greenwald, Poehlman, Uhlmann, & Banaji, 2009).



### Procedural Details

Considering the following procedural details might be helpful in creating IATs. First, the standard version of the IAT typically includes error feedback for incorrect responses (but see Olson & Fazio, 2004). In one variant, participants are presented with an error message that remains on the screen until they have pressed the correct key. In another variant, the error message replaces the stimulus for some time (e.g., 1,000 msec) before the next trial starts (participants are not required to correct their response in this variant). Both variants may have advantages and disadvantages, and it seems difficult to assess which variant may be superior. Regardless of which variant is adopted, it is important to note that Greenwald, Nosek, and Banaji (2003) make different recommendations for the treatment of errors in the scoring of IAT data, such that analyses of data obtained with the second, but not the first, variant should include an error penalty for incorrect responses (see below).

The length of intertrial intervals has been shown to have no significant effect on IAT scores (Greenwald & Nosek, 2001). In our own research, we typically use an intertrial interval of 250 msec, which seems long enough not to overwhelm participants but short enough to encourage quick responses. In addition, to display the particular key assignments on the screen throughout the task is strongly recommended, because the combined blocks can be very confusing if the key assignment is not available. With regard to task length, Greenwald et al. (2003) provided explicit recommendations on the number of trials in each of the five blocks (see Table 3.1). The two combined blocks—which include the critical trials of the task—are typically divided into two stages, with the initial stage of each block including a total of 20 trials and the subsequent stage, a total of 40 trials. The first two blocks typically include a total of 20 trials each; the fourth block usually includes a total of 40 trials. The recommendation to double the number of trials in the fourth block (compared to the 20 trials in the second block) is based on attempts to reduce the impact of block-order effects in the IAT (Nosek, Greenwald, & Banaji, 2005). Previous research has shown that the particular order of the two combined blocks can influence the size of IAT scores, such that they tend to be smaller when the third block includes association-incongruent trials followed by associations-congruent trials in the fifth block than when the third block includes association-congruent trials followed by associations-incongruent trials in the fifth block (Greenwald & Nosek, 2001). These block-related differences can be reduced by increasing the number of trials in the fourth block. Nevertheless, it is important to note that the IAT block-order structure is responsible for a number of dysfunctional, method-related effects (e.g., recoding of responses; for a more detailed discussion, see Teige-Mocigemba, Klauer, & Sherman, 2010), which inspired the development of IAT variants that avoid blocked presentations of association-congruent and association-incongruent trials (e.g., Rothermund et al., 2009; Teige-Mocigemba et al., 2008).

Another issue in this context concerns the recommendation to counterbalance the order of the association-congruent and the association-incongruent block in order to control for undesired block-order effects. Even though this recommendation seems quite plausible, it is important to note that counterbalancing can distort the rank order of individual differences in the relevant construct by introducing systematic error variance stemming from block-order effects (Gawronski, 2002). In fact, administering the same task with different orders implies that the resulting scores are not comparable across individuals, just as two scores on a self-report measure are not comparable if the items are presented in different orders. For this reason, we recommend keeping the block order constant (instead of counterbalancing), particularly if the research question involves individual differences.

Overall, previous studies using the IAT have reported satisfying reliability estimates, typically in the range between .70 and .90 (Cronbach's alpha; for a meta-analysis, see Hofmann et al., 2005); however, with regard to the use of multiple IATs in the same session, in our experience, the second and all subsequent IATs often show lower reliability estimates (e.g., Gawronski, 2002), sometimes going down to values between .50 and .60 (Cronbach's alpha). This issue puts some constraints on the use of multiple IATs in the same session, and it may partly explain why IAT scores tend to become lower as a function of enhanced practice with the task (Greenwald & Nosek, 2001).

### *Scoring*

IAT data are analyzed by comparing participants' performance on the two combined blocks, usually by calculating a difference score. In their original presentation of the IAT, Greenwald et al. (1998) suggested recoding the latencies of all trials with responses lower than 300 msec to 300 msec and latencies higher than 3,000 msec to 3,000 msec. Error trials should be excluded from analyses. For the use of analyses of variance (ANOVAs),



Greenwald et al. further recommended log-transforming the raw latencies before calculating the mean latencies in the two critical blocks. The so-called "IAT effect" is then calculated by subtracting the mean (log-transformed) latency of the third (fifth) block from the mean (log-transformed) latency of the fifth (third) block.

More recently, Greenwald et al. (2003) recommended a revised scoring algorithm that controls for individual variations in response latencies. In addition, the revised scoring algorithm includes a number of other steps that have the goal of increasing reliability. If the IAT is set up such that participants have to correct their responses on erroneous trials (see earlier discussion), the new algorithm includes the following seven steps: (1) Delete trials with latencies greater than 10,000 msec; (2) exclude participants for whom more than 10% of all trials have a latency of less than 300 msec; (3) compute the mean latencies in Stages 3, 4, 6, and 7; (4) compute the *inclusive* (not pooled) standard deviation[4] of a participant's latencies in Stages 3 and 6, and do the same for Stages 4 and 7; (5) compute the differences between the mean latencies in Stages 6 and 3 and between the mean latencies in Stages 7 and 4; (6) divide each difference score by the corresponding standard deviation calculated in step 3; (7) calculate the average of the two resulting scores. This algorithm is typically referred to as the D-algorithm. If the IAT is set up so that participants get error feedback but do not have to correct their responses on erroneous trials (see earlier discussion), the algorithm recommended by Greenwald et al. is identical to that outlined earlier, the only difference being that (1) steps 3 and 4 are completed on the basis of correct trials only, and (2) latencies on error trials are replaced by the mean value of all correct trials plus an error penalty of 600 msec before completing step 5. This algorithm is typically referred to as the D-600 algorithm.

## Derivates of the IAT

### *Single-Category IAT*

As outlined earlier, the task structure of the IAT is inherently comparative. For instance, scores revealed by a race IAT have to be interpreted as reflecting preferences for white over black individuals (or the other way round) rather than an absolute evaluation of blacks. To overcome this limitation, Karpinski and Steinman (2006) presented a modified variant of the

[4]The description in Table 4 of Greenwald et al. (2003) misleadingly describes this standard deviation as "pooled" (Anthony Greenwald, personal communication, April 7, 2009). The procedure outlined in the text (p. 201) correctly describes the standard deviation as "inclusive." Note, however, that a division by the inclusive standard deviation may overcorrect for systematic construct variance, because the inclusive (but not the pooled) standard deviation increases as a function of association-related differences between the congruent and the incongruent blocks.

IAT that assesses evaluations of single objects. Their Single-Category IAT (SC-IAT) uses the basic structure of the IAT; however, deviating from the standard IAT, the SC-IAT includes only one rather than two categories for the target dimension. In addition, the SC-IAT does not include any practice blocks, but simply uses the two combined blocks. For instance, in a SC-IAT application to assess evaluations of Pepsi, Karpinski and Steinman showed participants positive and negative words and pictures of Pepsi. In one block of the task, participants had to respond to negative words with one key, and to pictures of Pepsi and positive words with another key. In another block, the key assignment for Pepsi pictures was switched, such that participants had to respond to negative words and pictures of Pepsi with one key, and to positive words with another key. To avoid response biases as a result of disproportional left-hand and right-hand responses, words and pictures were not presented with equal frequencies, but in ratios that were closer to a balanced proportion of 50% for each response key. Note, however, that this strategy may still be suboptimal to completely eliminate response biases. For instance, in one of their four studies, Karpinski and Steinman had to remove close to 30% of their participants because of excessive error rates, which may have stemmed from task-induced response biases. To avoid such problems, other researchers preferred to adjust the proportion of positive and negative words, so that the proportion of left-hand and right-hand response can be kept at 50:50 (e.g., Galdi, Arcuri, & Gawronski, 2008); however, even though this strategy may reduce the likelihood of response biases, it also seems suboptimal because it undermines a counterbalanced presentation of positive and negative words. Finally, with regard to the use of pictures and words, it is worth noting that exclusive use of pictures for one dimension and words for the other dimension may also result in response biases, such that participants simplify their responses by recoding the task. For instance, in the previous example, participants may have recoded the task as "categorize the words as positive versus negative, and press the right (left) key each time a picture appears." To avoid such recoding effects, some researchers suggested including words and pictures for both response dimensions (e.g., Hofmann et al., 2007). Alternatively, one might limit the stimulus materials to either words or pictures for both dimensions, which could possibly resolve the problem of recoding. Another potential problem is that the SC-IAT retained the original block-order structure of the standard IAT, which has been shown to be associated with a number of dysfunctional effects (Rothermund et al., 2009; Teige-Mocigemba et al., 2008). Notwithstanding these issues, the SC-IAT has demonstrated satisfactory reliability estimates in the range of .70 and .90 (Cronbach's alpha), and even though the number of studies using the SC-IAT is significantly lower compared to the number of studies using the original IAT, there is assuring evidence for the validity of the SC-IAT.



### Single-Attribute IAT

Similar to the inspiration for the SC-IAT, there are cases where an attribute of interest does not have a "natural" counterpart that can be used as a contrast category in the standard IAT. To overcome this limitation, Penke et al. (2006) presented a Single-Attribute IAT (SA-IAT) that is structurally similar to Karpinski and Steinman's (2006) SC-IAT, the primary difference being the use of a single attribute rather than a single target object. The SA-IAT starts with a practice block for the target discrimination, followed by two combined blocks in which target trials are intermixed with attribute trials. For example, in an application of the SA-IAT to investigate *sociosexuality* (i.e., the tendency to engage in uncommitted sex), Penke et al. (2006) used the concepts *stranger* and *partner* as target categories and the concept *sex* as attribute. In a first block, participants had to practice the categorization of words related to *stranger* and *partner.* In the following combined blocks, the presentation of stranger-related and partner-related words was intermixed with presentations of sex-related words. In the first block, participants' task was to press one key when they saw a stranger-related word and another key when they saw a partner-related or a sex-related word. In the second block, participants were asked to press one key when they saw a stranger-related and a sex-related word and another key when they saw a partner-related word. Even though the SA-IAT has demonstrated satisfactory reliability estimates between .70 and .80 (Cronbach's alpha) in Penke et al.'s studies, it is important to note that the problems mentioned for the SC-IAT also apply to the SA-IAT. In addition, it is worth mentioning that we are not aware of any applications of the SA-IAT other than the original demonstration by Penke et al., which implies that current knowledge about the overall usefulness of the task is still limited.

### Personalized IAT

Olson and Fazio (2004) argued that the evaluative dimension in the standard IAT is often ambiguous to participants, in that it may refer either to one's own preferences or to preferences held by other people. For instance, the pleasant–unpleasant dimension in the race IAT could be interpreted as referring to what oneself finds pleasant or unpleasant or what other people find pleasant or unpleasant. As such, RI effects resulting from the categorization of the attribute words may be rooted in different representations, which Olson and Fazio refer to as *personal* versus *extrapersonal associations.* To overcome this limitation, they suggested a personalized variant of the IAT, in which the description of the evaluative attribute dimension is replaced by the labels "I like" and "I dislike." In addition, the error feedback is dropped from the task given that subjective judgments of liking versus disliking do not have a normatively correct answer. Even though Olson and Fazio's distinction between personal and extrapersonal associations has been the subject of conceptual (e.g., Gawronski, Peters, & LeBel, 2008) and empirical (e.g., Nosek & Hansen, 2008) controversies, the usefulness of the personalized IAT has been demonstrated in several studies showing that its outcomes differ from the ones obtained with the standard IAT in a predictable manner (e.g., Han, Olson, & Fazio, 2006). Further attesting to the personalized IAT's usefulness as an indirect measure, several studies have demonstrated satisfactory reliability estimates in the range of .70 to .90 (Cronbach's alpha); however, in many regards the basic structure of the personalized IAT can be criticized for the same methodological problems as the standard IAT (e.g., block-order structure, comparative nature). In addition, due to the lack of error feedback, it cannot be ruled out that respondents use the "I like" and "I dislike" answer categories to respond to items of the object dimension, thereby introducing an explicit component to the measure (Nosek & Hansen, 2008).

### *Single-Block IAT, Recoding-Free IAT*

A central problem with the block-order structure of the IAT is that it can produce method-related variance in IAT scores (Rothermund et al., 2009; Teige-Mocigemba et al., 2008) and undesired spillover effects on subsequent tasks (Klauer & Mierke, 2005). To overcome these problems, researchers developed IAT variants that include association-congruent and association-incongruent trials in a single, combined block. In Teige-Mocigemba et al.'s (2008) Single-Block IAT (SB-IAT), the particular key assignment for the target stimuli depends on whether the stimuli are presented in the upper or lower half of the screen; the key assignment for the attribute stimuli is the same for both the upper and lower half of the screen. For example, to assess preferences for white over black individuals, the first practice block of the SB-IAT may include presentations of black and white faces in the upper half of the screen, with black faces requiring a left-hand response and white faces a right-hand response. The second practice block would be identical to the first one, the only difference being that the faces are presented in the lower half of the screen, with the key assignment reversed. The third practice block combines the first two tasks, such that black and white faces are randomly presented either in the upper or lower half of the screen. Participants' task is to categorize the faces as black and white using the corresponding key assignments of the first two practice blocks. The fourth practice block includes the presentation of positive and negative words in the upper or lower half of the screen that have to be categorized with the same key assignment irrespective of their location. Finally, the critical fifth block combines target and attribution trials in alternating order, such that participants are presented with black and white faces and positive and negative words that appear in either the upper or lower half of the screen. The key assignment for the positive and negative words is the same regardless of whether they appear in the upper or lower half; however, black and white faces require responses in line with the key assignments trained in the practice blocks, such that they have to be categorized with one key mapping when they appear in upper half, but with the opposite mapping when they appear in the lower half.



Another IAT variant that aims at resolving the problems related to the IAT's block-order structure is Rothermund et al.'s (2009) Recoding-Free IAT (IAT-RF). The basic idea of the IAT-RF is very similar to that of the SB-IAT, such that the key assignments for the presented stimuli change randomly from trial to trial; however, differing from the SB-IAT, all stimuli appear in the center of the screen with display of the particular key assignments being shown before each individual trial.

The advantage of these two IAT variants is that they resolve various problems related to the block-order structure of the standard IAT, including different interpretations of the key labels in the two critical blocks as referring to one's own preferences versus preferences of other people (see Olson & Fazio, 2004), as well as undesired block-order effects on IAT scores (Greenwald & Nosek, 2001) and subsequent tasks (Klauer & Mierke, 2005). In Teige-Mocigemba et al.'s (2008) study, the SB-IAT demonstrated a somewhat wider range of reliability estimates compared to the standard IAT, with estimates ranging between .60 and .90 (Cronbach's alpha). Similarly, the reliability of IAT-RF tends to be somewhat lower compared to the standard IAT, with split-half coefficients ranging between .57 and .63. A potential reason for these differences might be that trial-by-trial switches of the key assignments could make the two IAT variants more cognitively demanding, thereby introducing more error noise. Alternatively, it seems possible that the reliability of the standard IAT is overestimated due to the contribution of systematic method variance to its internal consistency (e.g., Cronbach's alpha), which might therefore be reduced when the block-order structure is removed from the task.

### *Brief IAT*

Another recent variation of the standard IAT is Sriram and Greenwald's (2009) Brief IAT (B-IAT). The main inspiration for the B-IAT was to reduce the overall length of the standard IAT, while maintaining its psychometric properties. In line with the standard IAT, the B-IAT includes stimuli related to two dichotomous pairs of categories (e.g., black–white; pleasant–unpleasant); however, deviating from the standard IAT, participants are asked to focus on only two of the four categories. For this purpose, the task instructions in the B-IAT include a presentation of the stimuli of two focal categories (e.g., black faces, positive words), and participants are asked to memorize these stimuli for the following task. For the actual task, participants are instructed to press a "focal" response key each time they see one of the stimuli presented in the instructions, and to press a "nonfocal" response key for all other stimuli. The task itself includes two blocks.

For one of the two blocks, one of the two target categories is used as the focal category (e.g., black faces); for the other block, the alternative target category is used as the focal category (e.g., white faces). The focal status of the attribute category (e.g., positive words) is held constant across the two blocks. For example, in a B-IAT to assess preferences for white over black individuals, the first block may use black faces and positive words as the focal categories, whereas the second block uses white faces and positive words as the focal categories. The two blocks typically include a total of 40 trials each, divided in subblocks of 20 trials each. The resulting subblocks are administered such that participants first complete 20 trials with one focal set, then 20 trials with the alternative focal set, which is repeated in identical order in the second half of the task. Using Greenwald et al.'s (2003) D-algorithm, differences in participants' performance in the two focal sets are then interpreted as reflecting preferences for one over the other target category, in this case preferences for whites over blacks (or the other way round). Overall, the B-IAT demonstrated comparable performance to the standard IAT in Sriram and Greenwald's (2009) studies, even though the B-IAT showed satisfying validity only for instructional sets that used *pleasant* (rather than *unpleasant*) as the focal attribute category in evaluative B-IATs, and *self* (rather than *other*) as the focal target category in identity B-IATs. The range of reliability estimates was somewhat wider compared to the ones reported for the standard IAT, with estimates ranging between .55 and .95 (Cronbach's alpha). A potential problem with the B-IAT is that it retains many of the structural problems criticized about the standard IAT (e.g., block order). In addition, it is worth pointing out that the B-IAT is still relatively novel, and that additional research is needed to corroborate its validity.

## Go/No-Go Association Task

### *Task Structure*

Nosek and Banaji's (2001) GNAT, inspired by the basic logic of the standard IAT (Greenwald et al., 1998), is an attempt to make the task amenable for an assessment of associations related to a single concept rather than relative associations of two concepts. In a nutshell, the GNAT uses a go/no-go task in which participants are asked to show a "go" response to different kinds of target stimuli (e.g., by pressing the space bar) and a "no-go" response to distracter stimuli (i.e., no button press). For instance, in an adaptation of the GNAT to assess evaluative associations regarding black people, the task may include pictures of black people as targets, and pictures of nonblack people as distracters. In addition, the task may include positive and negative words, with positive words used as targets in one block of the task, and negative words as targets in a second block; that is, participants may be asked to show a "go" response to black faces and negative words, and a "no-go" response to nonblack faces and positive words in a first block of the task; in the second block, participants may be asked to show a "go" response to black faces and positive words, and a "no-go" response to nonblack faces and negative words. Differing from the latency-based algorithms for the IAT, GNAT scores are typically analyzed on the basis of error rates (see below). Thus, to increase the data basis for analyses of error data, the standard version of the GNAT typically includes a response deadline, such that participants are asked to show a "go" response to the targets before the expiration of that deadline (e.g., 600 msec).



### *Flexibility*

Like the IAT, the GNAT is very flexible in its application, in that targets and distracters may include a large variety of concepts and attributes. Previous research has applied the GNAT in various domains, including attitudes to social groups (e.g., Nosek & Banaji, 2001), self-related associations (e.g., Boldero, Rawlings, & Haslam, 2007), and associations to anxiety-provoking or phobic stimuli (e.g., Teachman, 2007).

### *Procedural Details*

In their original presentation of the GNAT, Nosek and Banaji (2001) investigated the impact of various procedural features, including different response deadlines (from 500 to 1,000 msec), different ratios of target and distracter trials (1:1 vs. 4:3), length of intertrial intervals (from 150 to 550 msec), and number of critical trials in the combined blocks. Overall, results were largely consistent across these variations. Nevertheless, most researchers employ similar task parameters, including response deadlines of 600–660 msec, equal ratios of target and distracter trials, intertrial intervals of 300 msec, and approximately 60 trials in each of the two critical blocks. Unfortunately, we did not find a lot of studies reporting reliability estimates for the GNAT. The few split-half reliability coefficients we were able to obtain from the literature were in the range between .45 and .75, with an (unweighted) average of .61. Even though the GNAT allows calculations of noncomparative scores, it again involves a block-order structure similar to the IAT, which has been linked to a number of dysfunctional method-related effects (see Teige-Mocigemba et al., 2010).

### *Scoring*

GNAT scores are usually calculated on the basis of error rates instead of response latencies. For this purpose, Nosek and Banaji (2001) suggested the use of signal detection theory (Green & Swets, 1966; Macmillan &

Creelman, 2006), such that differences in sensitivity scores ($d'$) between the two pairings of "go" trials (e.g., black-positive vs. black-negative) are interpreted as an index of associations between the concept of interest and the respective attributes. Sensitivity scores are calculated by (1) converting the proportion of hits (i.e., correct "go" responses to target items) and false alarms (i.e., incorrect "go" responses to distracter items) into standardized Z-scores and (2) subtracting the standardized proportion of false alarms from the standardized proportion of hits. An individual sensitivity score of zero indicates chance responding, such that participants were either unable to discriminate correctly between the stimuli or did not follow the instructions. For this reason, participants with scores equal (or close to zero) are typically excluded from analyses.

## Extrinsic Affective Simon Task

### *Task Structure*

Another task designed to overcome the problem of comparative assessments is De Houwer's (2003) EAST, which also avoids the block-order problems involved in many other tasks. In the critical block of the task, participants are presented with words depicting a target object of interest (e.g., the name of a beverage). The target words are presented in two different colors (e.g., yellow vs. blue), and participants' task is to respond to the target words with either a left-hand or a right-hand key, depending on the color in which the word is presented. Alternating with the presentation of colored target words, participants are presented with normatively positive and negative words in white ink color, which have to be categorized as positive or negative. Thus, the ultimate task for participants is to categorize the presented words in terms of their valence when they are presented in white ink color (i.e., positive vs. negative), and to categorize them in terms of their ink color when they are colored (e.g., blue vs. yellow). For instance, in an EAST application to assess evaluations of beer, participants may be presented with positive and negative words in white ink color and with the word *beer* in yellow ink color on some trials and in blue ink color on others. Participants' task is to press a left-hand key when they see a white word of negative valence or a word printed in blue ink, and to press a right-hand key when they see a white word of positive valence or a word printed in yellow. To the degree that participants show faster (or more accurate) responses to a colored word when the required response to this word is combined with a positive compared to a negative response, it is inferred that participants have positive associations with the object depicted by the colored word.

Overall, a typical EAST includes a total of three blocks, two practice blocks and one critical block. In the first block, participants are presented with the colored target words, which have to be categorized in terms of their ink color. In the second block, participants are presented with positive and negative words in white ink color, which have to be categorized in terms of their valence. In the critical third block, the two categorization tasks are combined, such that participants are presented with white and colored words in alternating order. Participants' task is to categorize the words in terms of their valence when they are presented in white ink, and to categorize the words in terms of ink color if they are colored.



In the standard version of the EAST, participants do not have to process the meaning of the colored target stimuli for the color-based responses required by the task. De Houwer and De Bruycker (2007a) speculated that this feature may be dysfunctional for a reliable assessment of mental associations (see De Houwer & De Bruycker, 2007b). To overcome this limitation, De Houwer and De Bruycker (2007a) presented a modified version of the EAST, in which participants are required to process the meaning of the target stimuli. The Identification EAST (ID-EAST) includes presentations of target and attribute words in uppercase and lower case. Positive and negative attribute words have to be categorized in terms of their valence irrespective of whether they are displayed in uppercase or lowercase. In contrast, the target words have to be categorized depending on whether they are presented in uppercase or lowercase. For instance, in an ID-EAST application to evaluations of beer, participants were randomly presented with positive and negative words and the word *beer* in either uppercase or lowercase. Participants were instructed to categorize the words in terms of their valence by pressing one of two response keys; however, for the word *beer*, participants were instructed to press one response key when the word was presented in uppercase and the opposite key when it was presented in lowercase. Because the attribute words were also presented in uppercase and lowercase, participants were therefore required to process the semantic meaning of the word *beer* before they were able to identify the correct response key.

### *Flexibility*

Even though the EAST was originally designed as a measure of evaluative responses, a number of studies have applied the task to other domains, such as the assessment of self-related associations. For instance, using the original variant of the EAST, Teige, Schnabel, Banse, and Asendorpf (2004) replaced the valence dimension of the white words with a self–other dimension that involved words related to self (e.g., *me*) and words related to others (e.g., *they*). As colored words, Teige et al. (2004) used adjectives that described three different traits (i.e., shyness, anxiousness, angriness). Similar adaptations seem possible for any kind of bipolar dimensions that can be used as attribute dimension. The use of targets is even more flexible, because the EAST is amenable to the use of single words representing a particular object or multiple words representing a category or even multiple dimensions (e.g., Teige et al., 2004; but see Gast & Rothermund, in press); however, it should be mentioned that the reliability of the EAST in Teige et al.'s (2004) study was rather low, with values between .19 and .24 (Cronbach's alpha), suggesting that the task might be suboptimal for the assessment of self-related associations or simultaneous assessment of multiple different associations.

### *Procedural Details*

In the original presentation of the EAST, De Houwer (2003) used 20 trials for each of the two practice blocks and a total of 120 trials in the critical block, divided into four subblocks of 30 trials each. The intertrial interval was set to 1,500 msec. The majority of subsequent EAST applications used between 30 and 40 trials for the two practice blocks and 120 trials for the critical block. In addition, most of these applications included the presentation of a fixation cross for 500 msec before each stimulus presentation, with an intertrial interval of 1,200 msec (excluding the fixation cross). The majority of reliability estimates we were able to find for EAST applications ranged between .40 and .50 (Cronbach's alpha), with an (unweighted) average score of .41 and reliability estimates ranging between .12 and .68.

### *Scoring*

In the original presentation of the EAST, De Houwer (2003) employed the traditional IAT algorithm proposed by Greenwald et al. (1998). Response latencies lower than 300 msec were recoded to 300 msec, and latencies higher than 3,000 msec were recoded to 3,000 msec, after which all latencies were log-transformed. Error trials were excluded from analyses. Valence scores were calculated by subtracting the mean log-transformed latency of responses to a colored word when its color was mapped with a "positive" key response from the mean log-transformed latency of responses to the same word when its color was mapped with a "negative" response. More recently, researchers adopted scoring procedures similar to Greenwald et al.'s (2003) D-600 algorithm for the analysis of EAST data (e.g., De Houwer & De Bruycker, 2007a; Schmukle & Egloff, 2006). The details of this algorithm are provided in the section on the standard IAT.

## Approach–Avoidance Tasks

### *Task Structure*

Another popular class of RI tasks can be broadly described as approach–avoidance tasks. The basic logic underlying approach–avoidance tasks is that approach reactions to a stimulus should be facilitated when the stimulus is positive rather than negative, whereas avoidance reactions should be facilitated when the stimulus is negative rather than positive. In the first published demonstration of such effects, Solarz (1960) found that participants were faster in pulling a lever toward them (approach) in response to positive compared to negative words. Conversely, participants were faster in pushing a lever away from them (avoidance) in response to negative compared to positive words. Expanding on these findings, Chen and Bargh (1999) showed that these effects emerge regardless of whether approach–avoidance responses are mapped with valence as the relevant stimulus feature (e.g., positive-approach, negative-avoidance vs. negative-approach, positive-avoidance) or with a valence-irrelevant feature (e.g., approach or avoidance as soon as the stimulus appears on the screen, with the onset of the stimulus presentation varying between 2 and 7 seconds). Irrespective of the particular task structure (i.e., valence-relevant vs. valence-irrelevant), approach–avoidance tasks can be described in terms of RI processes, such that the valence of a stimulus elicits a prepotent tendency to perform either an approach or an avoidance reaction that may either facilitate or interfere with the approach–avoidance response required in the task. Because approach–avoidance tasks can be designed to involve randomized presentations of association-congruent and association-incongruent trials, they are amenable to the assessment of individual associations without the dysfunctional effects of blocked presentations; however, such randomized presentations are not feasible if the responses in the task are based on the critical evaluative features of the stimuli (e.g., skin color of faces as the response criterion in a task to assess approach–avoidance tendencies to black and white faces). As we outline below, the latter type of approach avoidance tasks typically show higher internal consistencies, which implies a trade-off decision between the dysfunctional effects of blocked presentations and potentially low reliabilities.



### *Flexibility*

Approach–avoidance paradigms are quite flexible, in that they can be used for target stimuli of almost any kind (e.g., words, pictures) for both individual stimuli (e.g., a particular person) and general categories (e.g., black faces). Despite this flexibility in the use of target stimuli, it is important to note that the interpretation of RI effects in approach–avoidance tasks is still a subject of ongoing debates. Specifically, there are different views as to whether RI effects in approach–avoidance tasks are driven by (1) motivational orientations that are directly linked to particular motor actions, (2) motivational orientations that are related to processes of distance regulation, or (3) evaluative associations at the cognitive rather than motivational level. Each of these interpretations has unique implications for the suitability of approach–avoidance tasks for different types of research questions.

In line with a strong interpretation of embodied representations, early accounts attributed valence-related RI effects in approach–avoidance tasks to rigid links between motivational orientations and particular motor actions (e.g., Neumann, Förster, & Strack, 2003; see also Cacioppo, Priester, & Berntson, 1993). For instance, it has been argued (e.g., James, 1884) that contraction of the arm flexor muscle is inherently linked with an approach reaction (i.e., pulling an object toward the body), whereas contraction of the arm extensor muscle is inherently linked with an avoidance reaction (i.e., pushing an object away from the body). Based on these considerations, positive stimuli were assumed to elicit a spontaneous tendency for arm flexion, whereas negative stimuli were assumed to elicit a spontaneous tendency for arm extension. If these response tendencies are congruent with the response required in the task, quick and accurate responses should be facilitated. If, however, spontaneous approach–avoidance tendencies are incongruent with the ones required in the task, quick and accurate responses should be impaired.

Another interpretation emphasizes the role of distance regulation, arguing that RI effects in approach–avoidance tasks are driven by behavioral tendencies to increase the distance to aversive stimuli and to decrease the distance to appetitive stimuli (e.g., Seibt, Neumann, Nussinson, & Strack, 2008). This interpretation retains the notion of motivational orientations of the original embodiment account; however, it deviates from the embodiment interpretation, in that it does not assume rigid links between motivational orientations and particular motor actions. For instance, Markman and Brendl (2005) employed an approach–avoidance paradigm in which participants were presented with a picture of a corridor, with their first name appearing in the center of the screen. On each trial of the task, a positive or negative word appeared either above or below the name, which had to be moved either toward or away from the name with a corresponding joystick movement. Results showed that participants were faster in moving positive words toward their name and negative words away from their name compared to moving negative words toward the name and positive words away from the name. Importantly, these effects were independent of whether a particular movement required a contraction of the flexor or the extensor muscle, supporting the claim that RI effects in approach-avoidance tasks may be driven by motivational processes of distance regulation rather than inherent links between motivational orientations and motor actions. Similar findings were obtained by De Houwer, Crombez, Baeyens, and Hermans (2001), who used a picture of a mannequin to represent the self, which appeared either above or below a stimulus presented in the center of the screen. Depending on the particular conditions, participants were instructed to move the mannequin either toward or away from the stimulus. In line with Markman and Brendl's (2005) findings, results showed that participants were faster in moving the mannequin either toward positive stimuli and away from negative stimuli, irrespective of the required motor action.

Challenging the notion of motivational orientations implied by embodiment and distance regulation accounts, a third interpretation suggests that RI effects in approach–avoidance tasks depend on the positive or negative meaning assigned to a particular action in the description of the task (e.g., Eder & Klauer, 2009; Eder & Rothermund, 2008). This evaluative coding interpretation treats approach–avoidance tasks as functionally equivalent to other RI tasks, in that they are assumed to assess evaluative associations at a cognitive level rather than assess motivational orientations. In support of this claim, Eder and Rothermund found that participants were faster in moving a lever backward in response to positive words and faster in moving a lever forward in response to negative words when the required responses were described as *pull* (i.e., positive coding of backward movement) and *push* (i.e., negative coding of forward movement); however, these effects were reversed when the same movements were described as *upward* (i.e., positive coding of forward movement) and *downward* (i.e., negative coding of backward movement). These results suggest that RI effects in approach–avoidance tasks depend on the particular coding of the required motor actions as positive or negative, and their congruence with the valence of the stimulus. Expanding on these findings, Krieglmeyer, Deutsch, De Houwer, and De Raedt (2010) demonstrated that evaluative coding and distance regulation mechanisms may operate in parallel in approach–avoidance tasks. These authors experimentally dissociated effects of evaluative response labels and actual distance regulation, showing that stimulus valence interacted with both factors. Specifically, responses implying a decrease (increase) in the distance to positive (negative) stimuli were facilitated, over and above compatibility effects resulting from the valence of the response labels. Notwithstanding the parallel contribution of the two processes, the impact of coding-related mechanisms has important implications for accurate interpretations of measurement scores, because the particular descriptions of the required motor actions can influence the direction of RI effects in these tasks (i.e., whether a given RI effect reflects positive or negative associations). Thus, carefully designed instructions with proper response labels are important to avoid misinterpretations of the resulting scores.



### *Procedural Details*

In addition to the reviewed effects of task construal (e.g., Markman & Brendl, 2005) and evaluative coding (e.g., Eder & Rothermund, 2008), several procedural details are important to consider in studies using approach–avoidance tasks. One of these aspects concerns presentation time of the target stimuli. Even though the majority of studies employed supraliminal presentations of the target stimuli, a study by Alexopolus and Ric (2007) found reliable approach–avoidance effects even for subliminally presented stimuli. Another issue concerns the aforementioned difference between valence-relevant and valence-irrelevant response mappings. Challenging

Chen and Bargh's (1999) assumption that approach–avoidance actions are elicited even in the absence of an evaluative processing goal, Rotteveel and Phaf (2004) showed that RI effects in approach–avoidance tasks can disappear if participants have a nonevaluative processing goal. In their study, responses to happy and angry male and female faces showed clear evidence for emotion-related RI effects (i.e., approach facilitated by happy faces; avoidance facilitated by angry faces) when the required action was defined by the emotional expression of the target faces, but not when the required action was defined by the gender of the target face. Similar results were reported by Krieglmeyer and Deutsch (2010), who failed to observe reliable compatibility effects in Chen and Bargh's (1999) paradigm when participants responded to a stimulus feature other than valence.

Other important procedural features worth mentioning are the number of trials and intertrial intervals. In contrast to the IAT, there are no commonly accepted conventions for procedural details of approach–avoidance tasks. The reported intertrial intervals that we were able to obtain from the literature varied between 400 and 1,700 msec, with some studies including a fixation cross before the presentation of the target stimulus for a duration of 200–2,000 msec. For studies using randomly changing intertrial intervals (see earlier discussion), the ranges varied from 500–2,500 to 2,000–7,000 msec. The majority of studies used trial numbers around 100, with the number of trials ranging from 20 to 240.

Unfortunately, there are hardly any reports regarding the reliability of approach–avoidance tasks. To our knowledge, the only study that reports reliability estimates is a recent one by Krieglmeyer and Deutsch (2010) that systematically compared a joystick variant of Chen and Bargh's (1999) original paradigm, De Houwer et al.'s (2001) mannequin task, and a modified version of Chen and Bargh's (1999) paradigm using visual response feedback of increasing or decreasing stimulus size as a function of approach versus avoidance reactions (Rinck & Becker, 2007). Comparing reliability estimates across different outlier cutoffs, results showed satisfactory reliabilities for all three measures when the tasks required a valence-relevant categorization of the presented stimuli. The average reliability estimates were .86 for Chen and Bargh's (1999) original paradigm, .85 for De Houwer et al.'s (2001) mannequin task, and .72 for Rinck and Becker's (2007) modified paradigm with visual feedback (Spearman–Brown split-half coefficient); however, when the task involved valence-irrelevant categorizations of the target stimuli, reliability estimates dropped significantly. Under these conditions, average reliability estimates were .33 for Chen and Bargh's (1999) original paradigm, a negative score of -.18 for De Houwer et al.'s (2001) mannequin task, and .20 for Rinck and Becker's (2007) modified paradigm with visual feedback. These results support speculations that RI tasks may be more reliable when the valence-relevant dimension is task-relevant rather than task-irrelevant (e.g., De Houwer, 2009; Gawronski, Deutsch, et al., 2008).



### *Scoring*

In approach–avoidance tasks, the valence or motivational nature of a given stimulus (or stimulus category) is typically inferred by comparing the mean response latency of approach reactions to the mean latency of avoidance reactions given a particular stimulus (or stimulus category). To reduce the complexity of the obtained data, researchers often calculate a valence index, for instance, by subtracting the mean latency of approach reactions to a given stimulus (or stimulus category) from the mean latency of avoidance reactions to the same stimulus (or stimulus category). In this index, higher values indicate a more positive valence; however, in interpreting such difference scores, it is important to note that a value of zero does not reflect a neutral reference point, because baseline latencies for approach and avoidance reactions (e.g., pulling vs. pushing the lever of a joystick) tend to differ irrespective of stimulus valence. A similar problem applies to the interpretation of difference scores involving responses to different stimuli for the same type of response, because differences in the resulting index may be driven by contingent features of the stimuli (e.g., word length, word frequency) rather than differences in approach–avoidance reactions. For these reasons, we recommend calculating a different score between approach and avoidance reactions to the same stimulus (or stimulus category) and to interpret this score only in a relative (e.g., higher scores reflecting more positive valence) rather than absolute manner (e.g., scores higher than zero reflecting positive valence).

Another important issue in this context concerns the treatment of outliers. In contrast to the well-defined algorithms for the IAT, there are no commonly accepted guidelines for the treatment of outliers in approach–avoidance tasks. Nevertheless, a couple of issues should be taken into account when using approach–avoidance tasks. First, mean and median latencies differ significantly depending on whether the task involves a valence-relevant or valence-irrelevant matching. For instance, in Chen and Bargh's (1999) studies, the mean latency of approach–avoidance reactions was significantly higher when the task involved a valence-relevant matching (i.e., positive-approach, negative-avoidance vs. negative-approach, positive-avoidance) than when the task involved a valence-irrelevant matching (i.e., approach or avoidance as soon as the stimulus appears on the screen, with the onset of the stimulus presentation varying between 2 and 7 seconds). Thus, the same latency may reasonably be considered as invalid in one implementation, but valid in the other. For this reason, Chen and Bargh used a cutoff criterion of 4,000 msec in their first study that included an evaluative response dimension, but a cutoff of 1,500 in the second study that included a nonevaluative response dimension. As such, it seems useful to screen the distribution of response latencies to get a better sense of what should be considered a valid trial and where one should set a cutoff for the exclusion of outliers. A useful strategy to avoid arbitrary decisions is to include a response deadline that seems appropriate for the particular implementation of the task (e.g., 1,000 msec for word classifications) and to exclude the latencies from all trials on which participants did not respond within the deadline.

As for the lower end of the distribution, it again seems important to note that one and the same latency may represent an invalid anticipation for one implementation, but a valid trial for another. For instance, if participants' task is to respond as soon as a word appears on the screen (e.g., Chen & Bargh, 1999, Experiment 2), a latency of 200 msec may certainly represent a valid trial. If, however, participants' task is to show either an approach or an avoidance reaction depending on the valence of the presented word (e.g., Chen & Bargh, 1999, Experiment 1), it seems rather unlikely that participants actually identify the required response within 200 msec.

## Derivates of Approach–Avoidance Tasks

Approach–avoidance actions have also been adopted in two RI tasks that are structurally similar to the basic paradigm, yet sufficiently distinct to deserve separate consideration: the Evaluative Movement Assessment (EMA) by Brendl et al. (2005) and the Implicit Association Procedure (IAP) by Schnabel et al. (2006).

### *Evaluative Movement Assessment*

In Brendl et al.'s (2005) EMA, the meaning of a particular response (i.e., approach vs. avoidance) is manipulated by the visual presentation of movements on a computer screen. In this task, the participants' first names are randomly presented in a rectangular frame on either the left or the right side of a computer screen. On the other side of the screen, a letter string is displayed (e.g., *XXXXX*). After 700 msec, the letter string is replaced by a word. The participants' task is to move the presented word with the joystick as quickly as possible either toward or away from their names. In a first block of the task, participants are presented with normatively positive or negative distracter words. Participants have to move the presented word with the joystick toward their names when it is positive but away from their names when it is negative. In a second block, participants again have to move positive distracter words toward their names and negative distracter words away from their names. In contrast to the first block, however, positive and negative distracter words are interspersed with a set of target words, which represent the target objects whose valence the EMA is designed to assess (e.g., names of different beverages). These target words are typically presented to participants before the task with the instruction to memorize them. Depending on the particular block of the task, participants are required either to (1) treat



all target words as negative, such that they generally have to move them away from their names or (2) treat all target words as positive, such that they generally have to move them toward their names. The general effect in this paradigm is that participants are faster in moving positive target objects toward than away from their names, and faster in moving negative target objects away from than toward their names.

Even though the EMA is structurally similar to the standard approach–avoidance paradigm outlined earlier, several important differences should be taken into account when deciding which paradigm might be most useful for a particular research question. Aside from the use of horizontal instead of vertical movements, the most important difference is that the EMA has been designed to investigate *within*-participant differences in evaluation (i.e., personal rank orders of multiple-attitude objects) rather than *between*-participant differences (i.e., individual differences in the evaluation of a particular attitude object). Importantly, whereas the internal consistency of the EMA proved to be sufficiently high for within-participant differences in evaluation (Cronbach's alpha values around .80), internal consistency estimates for between-participant differences showed substantial variation with Cronbach's alpha values ranging from .32 to .73 (Brendl et al., 2005). To our knowledge, there are no published studies using the EMA aside from Brendl et al.'s original presentation, which implies that the available knowledge about the task is still limited.

### *Implicit Association Procedure*

A second adaptation of approach–avoidance actions is the IAP by Schnabel et al. (2006). Even though the IAP is structurally closer to the original paradigm than the EMA, it differs from earlier task variants, in that it assesses self-related rather than evaluative associations. For instance, in Schnabel et al.'s adaptation of the IAP to assess self-associations regarding shyness, participants were presented with self-related words (e.g., *my*, *me*) and other-related words (e.g., *your*, *them*) in a first block. Participants' task was to pull the lever toward them when the word on the screen was self-related but push the lever away when the word was other-related. In a second block, participants were again presented with self-related and other-related words; however, randomly interspersed with these words participants were additionally presented with words related to shyness (e.g., inhibited) and words related to non-shyness (e.g., *secure*). Participants' task was to pull the lever as quickly as possible for self-related and shyness-related words but to push the lever for other-related and non-shyness-related words. Finally, in a third block, participants were again presented with words related to self, other, shyness, and non-shyness; however, in contrast to the instructions for the second block, participants were now required to pull the lever as quickly as possible for self-related and non-shyness-related words but to push the lever for other-related and shyness-related words. The rationale underlying this task is that quick and accurate responses in the second block should be facilitated when participants have strong associations between themselves and shyness, whereas quick and accurate responses in the third block should be inhibited by strong associations between self and shyness. In Schnabel et al.'s study, the IAP showed convergent validity to the IAT and corresponding correlations to criterion measures. The reliability of the IAP seems satisfying, with Cronbach's alpha values in the range .75–.85 (e.g., Hogendoorn et al., 2008; Schnabel et al., 2006), though the number of published studies using the IAP is still too low to draw strong conclusion about the general validity of the task. In addition, it should be noted that IAP involves a block-order structure for association-congruent and association-incongruent trials that has been linked to a number of method-related problems (see Teige-Mocigemba et al., 2010).

## Action Interference Paradigm

### *Task Structure*

The AIP is structurally similar to the EAST, the only difference being that the meaning of the response keys is defined directly in the instructions rather than indirectly via categorizations of attribute-related stimuli. This difference makes the task suitable for very young children, who might get overwhelmed by the complex task requirements in the IAT, EAST, or other RI tasks. For instance, in one application to study the development of gender stereotypes, Banse et al. (2010) told young children that Santa Claus needed their help in delivering Christmas presents to other children. In a first block of the task, the children were told that the first family had a boy and a girl, and that the boy would like to get trucks and the girl would like to get dolls. The children were then shown pictures of trucks and dolls on the screen and were asked to give the presents to the kids as quickly as possible by pressing the buttons of a response box marked with pictures of the boy and the girl. In a second block, the children were told that they were now at the house of another family, which also had a boy and a girl; however, this boy would like to get dolls and the girl would like to get trucks. The children were then shown the same pictures of trucks and dolls, and they were asked to press the response buttons marked with the pictures of another boy and girl. Controlling for various procedural features, Banse et al. (2010) found that children were faster in making stereotype-congruent assignments (i.e., boy–truck, girl–doll) compared to stereotype-incongruent assignments (i.e., boy–doll, girl–truck), which was interpreted as evidence for spontaneous gender stereotyping in children.



### *Flexibility*

Among the RI tasks reviewed in this chapter, the AIP is the most content-specific measure, in that the original variant is particularly designed for the assessment of gender-stereotypical associations. Nevertheless, it seems possible to modify the AIP for the assessment of other types of associations. For instance, to assess evaluative associations in the domain of racial prejudice, the gender categories could be replaced by racial categories and the assignment task might involve the distribution of desirable and undesirable objects to black and white kids; however, it is important to point out that applications of the AIP to other domains require a different framing of the task in the instructions.

### *Procedural Details*

In a pilot study, Banse et al. (2010) investigated the potential impact of various procedural features in the AIP, including the order of stereotype-congruent versus stereotype-incongruent blocks and changes in the response set from the first to the second block (i.e., changing key mappings for girls and boys, with constant key mappings for trucks and dolls vs. constant key mappings for girls and boys, with changing key mappings for trucks and dolls). In addition to the significant main effect of stereotype congruency on latency scores, error rates were significantly higher in the second block after a change of the key assignment. Moreover, even though the expected RI effect was obtained in both of the two block-order conditions, this effect tended to be stronger when the stereotype-congruent block was administered before the stereotype-incongruent block. As such, the AIP suffers from the same block-order problems outlined earlier for the IAT (see Teige-Mocigemba et al., 2010).

Another important issue concerns the use of the AIP with older samples. Even though the simplicity of the task makes it amenable for studies involving very young children, it may be less suitable for studies involving older children or adults, who might be able to ignore the key descriptions in the task (e.g., left is girl, right is boy) and simply categorize the presented items in terms of their category membership (e.g., trucks go to the right, dolls go to the left). Given that such flexible reinterpretations of key labels seem more likely with increasing age, the advantageous features of the AIP for studies involving young children may turn out to be dysfunctional with studies involving older children and adults.

In their original presentation of the AIP, Banse et al. (2010) used 60 trials for each of the two blocks, divided in two subblocks of 30 trials each. To familiarize the children with the task, participants started with 10 practice trials, in which they had to categorize red and blue gift boxes in terms of their color. The intertrial interval was set to 1,000 msec. The internal consistencies of the AIP in the two studies by Banse et al. were .32 and .48, respectively (Cronbach's alpha).

### *Scoring*

The scoring procedure employed by Banse et al. (2010) was roughly similar to Greenwald et al.'s (2003) D-algorithm for the IAT. Latencies lower than 400 msec and higher than 10,000 msec were discarded. Error trials were excluded from analyses. An individual effect size index was calculated by subtracting a participant's mean response latency in the congruent block from the same participant's response latency in the incongruent block, which was then divided by the pooled standard deviation of that participant's response latencies in the two blocks.

## Interpretation of Measurement Scores

When we outlined the notion of RI, we argued that the reviewed measures assess automatic associations through the response-facilitating and response-interfering effects of prepotent response tendencies that presumably result from these associations. Even though this view is in line with common interpretations in the literature, a few issues should be taken into account when interpreting RI effects as an indicator of automatic associations.

### Multiple Processes Underlying RI Effects

An important issue in the interpretation of RI effects is that they are influenced not only by automatic associations but also by various other processes that contribute to participants' performance on the task. To illustrate this confound, consider a standard race IAT with the target categories "black" versus "white" and the attribute categories "pleasant" versus "unpleasant." In the combined blocks of this IAT, a black face may elicit a response tendency to press the "black" key and, to the degree that negative associations are activated, another response tendency to press the "unpleasant" key. If "black" and "negative" responses are mapped onto the same key (congruent block), responses will be facilitated. If, however, "black" and "negative" responses are mapped onto different keys (incongruent block), then the prepotent tendency to press the "negative" key has to be inhibited, so that the accurate tendency to press the "black" key can be executed. Importantly, because inhibition of the prepotent response tendency requires executive control processes, the impact of race-related associations is confounded with control processes in the traditional calculation of IAT scores.

To overcome this problem, researchers have developed a number of alternative procedures to analyze data obtained with RI tasks (e.g., Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007; Payne, 2008; Stahl & Degner, 2007). The main purpose of these procedures is to quantify the individual contributions of multiple distinct processes to task performance instead of relying on a single index of automatic associations. Because the mathematical underpinnings of these procedures are extensively discussed by Klauer, Stahl, and Voss (Chapter 12, this volume), we refrain from discussing them in more detail here. Nevertheless, we want to mention two important implications of process confounds involved in RI tasks. First, in studies using traditional measurement scores as independent variables (e.g., prediction of behavior), the obtained relations to a criterion measure may be driven by an overlap in the operation of executive control processes (e.g., reflecting individual differences in working memory capacity) rather than by the presumed influence of mental associations (see Klauer, Schmitz, Teige-Mocigemba, & Voss, 2010). Second, in studies using traditional measurement scores as dependent variables, the measurement scores may be influenced by experimentally created changes in executive control processes rather than genuine changes in mental associations (see Sherman et al., 2008). These ambiguities can be resolved by the use of more sophisticated mathematical procedures, such as the ones reviewed by Klauer et al. (Chapter 12, this volume).



### Absolute versus Relative Interpretations

Another important issue concerns absolute versus relative interpretations of the measurement scores revealed by RI tasks. Many of the reviewed scoring procedures involve difference scores that can produce values higher or lower than zero. This metric is often used to infer associations in one direction if the resulting score is higher than zero, and associations in the opposite direction if the score is lower than zero, with a value of zero being interpreted as a neutral reference point. Even though absolute interpretations of this kind are rather common in the literature, we consider them problematic for several reasons. Aside from the fact that the metric of any given measure remains ambiguous without proper calibration (see Blanton & Jaccard, 2006), contingent features of the employed stimulus materials have been shown to influence both the size and the direction of measurement scores revealed by RI tasks (e.g., Bluemke & Fiedler, 2009; Bluemke & Friese, 2006; Steffens & Plewe, 2001). Because it is virtually impossible to quantify the contribution of material effects to the overall size of RI effects, absolute interpretations of measurement scores do not seem justified, regardless of whether they involve characteristics of individual participants (e.g., participant X shows a preference for whites over blacks) or samples (e.g., 80% of the sample showed a preference for whites over blacks).

Even though these issues may seem rather devastating for the use of RI tasks, it is important to note that most research questions in personality and social psychology do not require absolute interpretations. In fact, a large majority of psychological studies is based on relative interpretations in the sense that measurement scores are compared across either different groups or different individuals. Whereas the first case involves experimental designs in which measurement scores are compared across different groups (e.g., participants in the experimental group show higher scores compared to participants in the control group), the latter case involves individual-difference designs in which measurement scores are compared across different participants (e.g., participant *A* has a higher score compared to participant *B*). Because both types of research designs are based on relative rather than absolute interpretations of measurement scores, the aforementioned problems of material effects do not necessarily undermine the usefulness of RI tasks in psychological research, at least as long as the particular research question does not require absolute interpretations of measurement scores.

## Automaticity and Control

Our discussion of multiple processes underlying RI effects already has suggested that standard data-analytic procedures confound the influence of automatic associations with different kinds of controlled processes (Conrey et al., 2005; Klauer et al., 2007; Payne, 2008; Stahl & Degner, 2007). Another important question in this context is whether the assessed associations per se are automatic. As we outlined at the beginning of this chapter, the automatic nature of these associations is not guaranteed by the indirect nature of the measurement procedures, but has to be tested empirically by means of carefully designed studies (De Houwer et al., 2009). Unfortunately, the available evidence in this regard is still scarce. Concerning the IAT, the available evidence suggests that the associations assessed by the IAT (1) can be controlled to some extent, even though participants have less control over their responses compared to traditional self-report measures; (2) can have unconscious origins, as well as unconscious effects, even though the associations per se may be consciously accessible; (3) are not necessarily unintentional, because intentional retrieval of information has been shown to influence IAT scores; and (4) are efficient, in that reduced processing capacity does not seem to reduce IAT scores. A review of relevant evidence in this regard is provided by De Houwer et al. for the IAT and evaluative priming tasks. Evidence for other RI tasks is still scarce. The only two exceptions of which we are aware are studies by Schnabel et al. (2006), demonstrating the robustness of the IAP against faking, and by Langner et al. (2010), showing that responses in approach–avoidance tasks can be faked.



Another important issue in this context concerns the use of RI tasks as dependent measures to investigate the presumed automatic nature of a process or experimental effect. Consider, for example, the well-replicated finding that enhanced salience of one's mortality enhances the extremity of worldviews, including prejudice against outgroups (for a review, see Pyszczynski, Greenberg, & Solomon, 1999). To test the automatic nature of this effect, one might be tempted to use an indirect measure of prejudice as the dependent measure and infer that mortality salience effects are automatic if the measure shows higher levels of prejudice under mortality salience compared to control conditions. Even though such inferences are relatively common in the literature, they are not justified, because RI tasks assess the *outcome* of the employed manipulation, not the *process* that leads to the outcome. To illustrate this issue, consider a study by Blair, Ma, and Lenton (2001), in which participants were asked to imagine vividly either a strong or a weak woman, and then completed an IAT designed to measure stereotypical associations between gender and strength. Results showed that IAT scores of gender-stereotypical associations were lower when participants were asked to imagine a strong woman than when they were asked to imagine a weak woman. Does this finding show that mental imagery is an automatic process? Obviously not, as the process of imagining a counterstereotypical woman was conscious, intentional, and certainly controllable in Blair et al.'s study. Of course, some experimental manipulations may influence the scores revealed by RI tasks through automatic processes. This, however, does not justify the reverse inference that RI tasks can be used to demonstrate the automatic nature of a given process or effect, which would be a prime example of the logical fallacy of *affirming the consequent*. In other words, the fact that automatic processes can lead to variations in the measurement scores revealed by RI tasks does not imply that any variation in measurement scores is driven by automatic processes.

## RI Tasks as Measures of Accessibility

If the question of automaticity is indeed as thorny as we argue, what exactly is the benefit of using RI tasks in psychological research? As we noted in our brief history of RI tasks, recent theorizing in this area has shifted toward interpretations that are closer to the notion of accessibility (e.g., Gawronski & Bodenhausen, 2006; Strack & Deutsch, 2004). According to these theories, RI tasks and other kinds of indirect measures assess the momentary accessibility of mental associations, irrespective of whether these associations are regarded as valid or invalid. The latter type of validity qualification is the central feature of traditional self-report measures, in which participants are typically asked to indicate their agreement or disagreement with a particular statement (e.g., "Please rate how much you agree with the statement ... "). The most important finding in this context is that accessible associations can influence overt behavior, even when these associations are rejected as invalid in standard self-report measures. This issue is most prominently reflected in the finding that indirect measures—such as the RI tasks reviewed in this chapter—predict behaviors that are difficult to predict with standard self-report measures (for reviews, see Friese et al., 2008; Perugini et al., 2010). Thus, even though claims about automatic features should be treated with caution, unless they are backed up with empirical data (De Houwer et al., 2009), RI tasks clearly represent a useful addition to the toolbox of cognitive instruments, providing a better understanding of the determinants of human behavior.

## Comparison of Measures

Given the large variety of available RI tasks, the ultimate question is which of them to select for one's own research. Our response to that question is: It depends. In our view, RI tasks are tools, and different types of research questions require different kinds of tools. Thus, instead of recommending a particular measure as the "best" one, we try to compare the reviewed tasks with regard to a few features that we deem important for different kinds of research questions.

A first issue concerns reliability, which varies considerably across tasks (see Table 3.2). Whereas some measures have consistently shown satisfying reliability estimates across different applications, others suffer from large variations or clearly unsatisfactory psychometric properties. Even though low reliability is a problem for any kind of research, this issue seems particularly important for studies investigating individual differences, which presuppose consistent rank orders of participants within a given measure (internal consistency).[5] Second, it is worth noting that there is considerable variation in the number of published studies using different kinds of RI tasks (see Table 3.2). Even though these variations may be due to multiple different factors, the number of published studies provides a good proxy of the amount of knowledge available about a given measure, as well as its success in producing publishable data (to be corrected by year of publication). Third, the reviewed RI tasks differ considerably with regard to their flexibility, in that some tasks can be used to assess different kinds of associations (e.g., semantic, evaluative), whereas others are more specific in the type of questions for which

[5] Unfortunately, it is still not common practice to report reliability estimates for indirect measures. It is, therefore, difficult to assess whether the published reliability estimates imply a reporting-related overestimation of the true reliability of some tasks (i.e., reliability may be reported when it is high, but not when it is low).

**TABLE 3.2. Overview of RTI Tasks**

| Task | Publication year | Number of publications | Reliability estimates | Types of associations | Target objects | Trial structure | Number of trials | Intertrial interval |
|---|---|---|---|---|---|---|---|---|
| IAT | 1998 | High | .70–.90 [a] | Evaluative, semantic | Pairs of targets | Blocked | ~200 | 250 msec |
| SC-IAT | 2006 | Moderate | .70–.90 | Evaluative, semantic | Individual targets | Blocked | ~200 | 150 msec |
| SA-IAT | 2006 | Low | .70–.80 | Evaluative, semantic | Pairs of targets | Blocked | ~360 | 150 msec |
| P-IAT | 2004 | Moderate | .70–.90 | Evaluative | Pairs of targets | Blocked | ~200 | 250 msec |
| SB-IAT | 2008 | Low | .60–.90 | Evaluative, semantic | Pairs of targets | Random | ~160 | 500 msec |
| IAT-RF | 2009 | Low | .55–.65 | Evaluative, semantic | Pairs of targets | Random | ~200 | 200 msec |
| B-IAT | 2009 | Low | .55–.95 | Evaluative, semantic | Pairs of targets | Blocked | ~80 | 400 msec |
| GNAT | 2001 | Moderate | .45–.75 | Evaluative, semantic | Individual targets | Blocked | ~120 | 300 msec |
| EAST | 2003 | Moderate | .15–.65 | Evaluative, semantic | Individual targets | Random | ~160 | 1,500 msec |
| ID-EAST | 2007 | Low | .60–.70 | Evaluative, semantic | Individual targets | Random | ~260 | 1,500 msec |

*(continued)*

TABLE 3.2. *(continued)*

| Task | Publication year | Number of publications | Reliability estimates | Types of associations | Target objects | Trial structure | Number of trials | Intertrial interval |
|---|---|---|---|---|---|---|---|---|
| AAT | 1999 | Moderate | .00–.35<br>.70–.90[b] | Evaluative | Individual targets | Blocked, random | ~100 | 1,000 msec<br>500–7,000 msec[c] |
| EMA | 2005 | Low | .30–.75<br>~.80[d] | Evaluative | Individual targets | Random | ~280 | 600 |
| IAP | 2006 | Low | .75–.85 | Self-related | Pairs of targets | Blocked | ~280 | 600 msec |
| AIP | 2010 | Low | .30–.50 | (Content-specific)[e] | Pairs of targets | Blocked | ~120 | 1,000 msec |

*Note.* This overview includes the publication year of each task's representative references; overall number of publications that have used the task; approximate range of reported reliability estimates; flexibility of applications regarding types of associations and types of target objects; structure for the presentation of association-congruent versus association-incongruent trials; prototypical number of trials; and prototypical intertrial intervals. IAT, Implicit Association Test; SC-IAT, Single-Category Implicit Association Test; SA-IAT, Single-Attribute Implicit Association Test; P-IAT, Personalized Implicit Association Test; SB-IAT, Single-Block Implicit Association Test; IAT-RF, Recoding-Free Implicit Association Test; B-IAT, Brief Implicit Association Test; GNAT, Go/No-Go Association Task; EAST, Extrinsic Affective Simon Task; ID-EAST, Identification Extrinsic Affective Simon Task; AAT, Approach-Avoidance Tasks; EMA, Evaluative Movement Assessment; IAP, Implicit Association Procedure; AIP, Action Interference Paradigm.

[a]Reliability estimates tend to be lower for second and subsequent IATs if more than one IAT is administered in the same session.

[b]Reliability estimates differ depending on whether approach–avoidance responses involve valence-relevant or valence-irrelevant categorizations, with valence-irrelevant categorizations showing lower reliability estimates compared to valence-relevant categorizations; numbers in the upper row show estimates for valence-irrelevant categorizations; numbers in the lower row show estimates for valence-relevant categorizations.

[c]Numbers in the upper row show the prototypical intertrial interval for tasks with constant intervals; numbers in the lower row show the range of intertrial intervals for tasks with randomly varying intervals.

[d]Reliability estimates differ depending on whether the scores involve within-participant comparisons of preferences for different objects or between-participant comparisons of evaluations of the same object, with between-participant comparisons showing lower reliability estimates compared to within-participant comparisons; numbers in the upper row show estimates for between-participant comparisons; numbers in the lower row show estimates for within-participant comparisons.

[e]Previous applications are limited to gender-stereotypical associations, though alternative applications seem possible.



they can be used (see Table 3.2). Finally, it is important to point out that none of the reviewed measures is perfect, and that any choice between these tasks involves a trade-off decision between desirable and undesirable features. Examples of these features include the block-order structure of the task, its applicability to individual versus pairs of target concepts, potential response biases due to asymmetric numbers of left-hand and right-hand responses, the overall length of the task, and its suitability for children. Of course, the relative importance of these features depends on the particular research question, which makes it difficult to make strong recommendations on a priori grounds. Nevertheless, we hope that our review is helpful in making informed decisions about which particular measure might be most useful for a given research question. For readers interested in learning more about method-related aspects of RI tasks and other kinds of indirect measures, we recommend the readings listed at the end of this chapter, as well as the method-related chapters in Gawronski and Payne's (2010) *Handbook of Implicit Social Cognition.*

Before we conclude, it is important to answer an additional question that have we deliberately avoided so far: the convergent validity of the reviewed measures. Unfortunately, the available evidence on this issue is still somewhat scarce. Even though there is preliminary evidence for convergent validity between some of the reviewed tasks (e.g., Cunningham, Preacher, & Banaji, 2001; De Houwer & De Bruycker, 2007b; Karpinski & Steinman, 2006; Neumann, Hülsenbeck, & Seibt, 2004; Nosek & Banaji, 2001; Schnabel et al., 2006), in many studies the overlap between indirect measures has been far from satisfactory (e.g., Gast & Rothermund, in press; Olson & Fazio, 2003; Teige et al., 2004). In our view, one of the most important studies in this context was conducted by Olson and Fazio (2003), who showed that race applications of Fazio et al.'s (1986) evaluative priming task and the IAT are significantly correlated with each other only when participants are instructed to pay attention to the race of the employed face primes in the evaluative priming task. Olson and Fazio (2003) interpreted this finding as evidence for their claim that sequential priming tasks assess evaluative associations pertaining to the individual faces used as prime stimuli (i.e., exemplar-related associations), whereas the IAT assesses evaluative associations pertaining to the categories that are used to classify the faces in the task (i.e., category-related associations). Note, however, that in Olson and Fazio's (2003) study the reliability of the evaluative priming task differed remarkably across the two attention conditions (split-half correlations of $r = .04$ vs. $r = .39$). Thus, the obtained differences in correlations to the IAT may be due to differences in the reliability of the evaluative priming task rather than in the type of associations assessed by the two measures (see Gawronski, Cunningham, LeBel, & Deutsch, 2010). In fact, the limited available evidence so far suggests that the reviewed RI tasks indeed show convergent validity, and that dissociations are most often due to low reliabilities of the employed tasks (e.g., Cunningham et al., 2001; De Houwer

& De Bruycker, 2007b); however, based on evidence for method-related dissociations between measures based on different underlying mechanisms (e.g., Deutsch & Gawronski, 2009; Gawronski & Bodenhausen, 2005), it seems imperative for future research to confirm the convergent validity of the various RI tasks and other indirect measures.

## Conclusion

Over the past years, indirect measures of automatic associations have enjoyed an enormous popularity within and beyond the boundaries of social psychology. This chapter has provided a method-oriented overview of a particular class of indirect measures, namely, measures based on RI. These tasks not only represent the majority of the currently available indirect measures but they also have provided significant insights into the mental processes underlying human behavior. We hope that this chapter serves as a useful resource for anyone who is interested in using RI tasks in their research.

## Recommended Reading

### General Readings

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

Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology, 50*, 141–150.

### Implicit Association Test

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: IV: What we know (so far) about the method.. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 59–102). New York: Guilford Press.

### Single-Category IAT

Karpinski, A., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology, 91*, 16–32.

### Single-Attribute IAT

Penke, L., Eichstaedt, J., & Asendorpf, J. B. (2006). Single Attribute Implicit Association Tests (SA-IAT) for the assessment of unipolar constructs: The case of sociosexuality. *Experimental Psychology, 53*, 283–291.



### Personalized IAT

Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extra-personal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology, 86*, 653–667.

### Single-Block IAT

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2008). Minimizing method-specific variance in the IAT: The Single Block IAT. *European Journal of Psychological Assessment, 24*, 237–245.

### Recoding-Free IAT

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

### Brief IAT

Sriram, N., & Greenwald, A. G. (2009). The Brief Implicit Association Test. *Experimental Psychology, 56*, 283–294.

### Go/No-Go Association Task

Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition, 19*, 625–666.

### Extrinsic Affective Simon Task

De Houwer, J. (2003). The Extrinsic Affective Simon task. *Experimental Psychology, 50*, 77–85.

### Identification EAST

De Houwer, J., & De Bruycker, E. (2007). The Identification-EAST as a valid measure of implicit attitudes toward alcohol-related stimuli. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 133–143.

### Approach–Avoidance Tasks

Krieglmeyer, R., & Deutsch, R. (2010). Comparing measures of approach–avoidance behavior: The mannequin task vs. two versions of the joystick task. *Cognition and Emotion, 24*, 810–828.

### Evaluative Movement Assessment

Brendl, C. M., Markman, A. B., & Messner, C. (2005). Indirectly measuring evaluations of several attitude objects in relation to a neutral reference point. *Journal of Experimental Social Psychology, 41*, 346–368.

### Implicit Association Procedure

Schnabel, K., Banse, R., & Asendorpf, J. B. (2006). Employing automatic approach and avoidance tendencies for the assessment of implicit personality self-concept: The Implicit Association Procedure (IAP). *Experimental Psychology, 53,* 69–76.

### Action Interference Paradigm

Banse, R., Gawronski, B., Rebetez, C., Gutt, H., & Morton, J. B. (2010). The development of spontaneous gender stereotyping in childhood: Relations to stereotype knowledge and stereotype flexibility. *Developmental Science, 13,* 298–306.

## References

Alexopolus, T., & Ric, F. (2007). The evaluation–behavior link: Direct and beyond valence. *Journal of Experimental Social Psychology, 43,* 1010–1016.

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

Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between explicit and implicit 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, III, J. S. Nairne, I. Neath, & A. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 117–150). Washington, DC: American Psychological Association.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science, 7,* 136–141.

Banse, R., Gawronski, B., Rebetez, C., Gutt, H., & Morton, J. B. (2010). The development of spontaneous gender stereotyping in childhood: Relations to stereotype knowledge and stereotype flexibility. *Developmental Science, 13,* 298–306.

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

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic activation with a pronunciation task. *Journal of Personality and Social Psychology, 32,* 104–128.

Blair, I. V., Ma, J., & Lenton, A. (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. (2006). Arbitrary metrics in psychology. *American Psychologist, 61,* 27–41.

Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the Implicit Association Test: Implications for criterion prediction. *Journal of Experimental Social Psychology, 42,* 192–212.



Bluemke, M., & Fiedler, K. (2009). Base rate effects on the IAT. *Consciousness and Cognition, 18,* 1029–1038.

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

Boldero, J. M., Rawlings, D., & Haslam, N. (2007). Convergence between GNAT-assessed implicit and explicit personality. *European Journal of Personality, 21,* 341–358.

Brendl, C. M., Markman, A. B., & Messner, C. (2005). Indirectly measuring evaluations of several attitude objects in relation to a neutral reference point. *Journal of Experimental Social Psychology, 41,* 346–368.

Cacioppo, J. T., Priester, J. R., & Berntson, G. G. (1993). Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology, 65,* 5–17.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25,* 215–224.

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

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24,* 349–354.

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). The Extrinsic Affective Simon task. *Experimental Psychology, 50,* 77–85.

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 procedural and functional level. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 361–390). New York: Psychology Press.

De Houwer, J., Crombez, G., Baeyens, F., & Hermans, D. (2001). On the generality of the affective Simon effect. *Cognition and Emotion, 15,* 189–206.

De Houwer, J., & De Bruycker, E. (2007a). The Identification-EAST as a valid measure of implicit attitudes toward alcohol-related stimuli. *Journal of Behavior Therapy and Experimental Psychiatry, 38,* 133–143.

De Houwer, J., & De Bruycker, E. (2007b). The Implicit Association Test outperforms the Extrinsic Affective Simon Task as an implicit measure of interindividual differences in attitudes. *British Journal of Social Psychology, 46,* 401–421.

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.

Eder, A. B., & Klauer, K. C. (2009). A common-coding account of the bi-directional evaluation–behavior link. *Journal of Experimental Psychology: General, 138,* 218–235.

Eder, A. B., & Rothermund, K. (2008). When do motor behaviors (mis)match affective stimuli?: An evaluative coding view of approach and avoidance reactions. *Journal of Experimental Psychology: General, 137,* 262–281.

Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition, 25,* 603–637.

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.

Förster, J., & Liberman, N. (2007). Knowledge activation. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2nd ed., pp. 201–231). New York: Guilford Press.

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.

Galdi, S., Arcuri, L., & Gawronski, B. (2008). Automatic mental associations predict future choices of undecided decision makers. *Science, 321,* 1100–1102.

Gast, A., & Rothermund, K. (in press). When old and frail is not the same: Dissociating category and stimulus effects in four implicit attitude measurement methods. *Quarterly Journal of Experimental Psychology.*

Gawronski, B. (2002). What does the Implicit Association Test measure?: A test of the convergent and discriminant validity of prejudice related IATs. *Experimental Psychology, 49,* 171–180.

Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology, 50,* 141–150.

Gawronski, B., & Bodenhausen, G. V. (2005). Accessibility effects on implicit social cognition: The role of knowledge activation versus retrieval experiences. *Journal of Personality and Social Psychology, 89,* 672–685.

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

Gawronski, B., Cunningham, W. A., LeBel, E. P., & Deutsch, R. (2010). Attentional influences on affective priming: Does categorization influence spontaneous evaluations of multiply categorizable objects? *Cognition and Emotion, 24,* 1008–1025.

Gawronski, B., Deutsch, R., LeBel, E. P., & Peters, K. R. (2008). Response interference as a mechanism underlying implicit measures: Some traps and gaps in the assessment of mental associations with experimental paradigms. *European Journal of Psychological Assessment, 24,* 218–225.

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



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

Gawronski, B., & Payne, B. K. (Eds.). (2010). *Handbook of implicit social cognition: Measurement, theory, and applications.* New York: Guilford Press.

Gawronski, B., Peters, K. R., & LeBel, E. P. (2008). What makes mental associations personal or extra-personal?: Conceptual issues in the methodological debate about implicit attitude measures. *Social and Personality Psychology Compass, 2,* 1002–1023.

Gawronski, B., & Sritharan, R. (2010). Formation, change, and contextualization of mental associations: Determinants and principles of variations in implicit measures. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 216–240). New York: Guilford Press.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

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

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109,* 3–25.

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

Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (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. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie, 48,* 85–93.

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., Poehlman, T. A., Uhlmann, E., & 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.

Han, H. A., Olson, M. A., & Fazio, R. H. (2006). The influence of experimentally created extrapersonal associations on the Implicit Association Test. *Journal of Experimental Social Psychology, 42,* 259–272.

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 measure. *Personality and Social Psychology Bulletin, 31,* 1369–1385.

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.

Hogendoorn, S. M., Wolters, L. H., Vervoort, L., Prins, P. J. M., Boer, F., & de

Haan, E. (2008). An indirect and direct measure of anxiety-related perceived control in children: The Implicit Association Procedure (IAP) and Anxiety Control Questionnaire for Children (ACQ-C). *Journal of Behavior Therapy and Experimental Psychiatry, 39*, 436–450.

James, W. (1884). What is an emotion? *Mind, 9*, 188–205.

Karpinski, A., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology, 91*, 16–32.

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 IAT effect. *Personality and Social Psychology Bulletin, 31*, 208–217.

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.

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

Krieglmeyer, R., & Deutsch, R. (2010). Comparing measures of approach–avoidance behavior: The manikin task vs. two versions of the joystick task. *Cognition and Emotion, 29*, 810–828.

Krieglmeyer, R., Deutsch, R., De Houwer, J., & De Raedt, R. (2010). Being moved: Valence activates approach–avoidance behavior independently of evaluation and approach–avoidance intentions. *Psychological Science, 21*, 607–613.

Langner, O., Ouwens, M., Muskens, M., Trumpf, J., Becker, E. S., & Rinck, M. (2010). Faking on direct, indirect, and behavioural measures of spider fear: Can you get away with it? *Cognition and Emotion, 24*, 549–558.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.

Maison, D., Greenwald, A. G., & Bruin, R. (2004). Predictive validity of the Implicit Association Test in studies of brands, consumer attitudes, and behavior. *Journal of Consumer Psychology, 14*, 405–415.

Markman, A. B., & Brendl, C. M. (2005). Constraining theories of embodied cognition. *Psychological Science, 16*, 6–10.

Moors, A., & De Houwer, J. (2006). Automaticity: A conceptual and theoretical analysis. *Psychological Bulletin, 132*, 297–326.

Neumann, R., Hülsenbeck, K., & Seibt, B. (2004). Attitudes toward people with AIDS and avoidance behavior: Automatic and reflective bases of behavior. *Journal of Experimental Social Psychology, 40*, 543–550.

Neumann, R., Förster, J., & Strack, F. (2003). Motor compatibility: The bidirectional link between behavior and emotion. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 371–391). Mahwah, NJ: Erlbaum.

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. (2001). The Go/No-Go Association Task. *Social Cognition, 19*, 625–666.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin, 31*, 166–180.



Nosek, B. A., & Hansen, J. (2008). Personalizing the Implicit Association Test increases explicit evaluation of target concepts. *European Journal of Psychological Assessment, 24*, 226–236.

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 extra-personal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology, 86*, 653–667.

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

Paulhus, D. L. (1984). Two component models of social desirable responding. *Journal of Personality and Social Psychology, 46*, 598–609.

Payne, B. K. (2008). What mistakes disclose: A process dissociation approach to automatic and controlled processes in social psychology. *Social and Personality Psychology Compass, 2*, 1073–1092.

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

Penke, L., Eichstaedt, J., & Asendorpf, J. B. (2006). Single Attribute Implicit Association Tests (SA-IAT) for the assessment of unipolar constructs: The case of sociosexuality. *Experimental Psychology, 53*, 283–291.

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

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

Petty, R. E., Tormala, Z. L., Briñol, P., & Jarvis, W. B. G. (2006). Implicit ambivalence from attitude change: An exploration of the PAST model. *Journal of Personality and Social Psychology, 90*, 21–41.

Pyszczynski, T., Greenberg, J., & Solomon, S. (1999). A dual process model of defense against conscious and unconscious death-related thought: An extension of terror management theory. *Psychological Review, 106*, 835–845.

Richetin, J., Perugini, M., Adjali, I., & Hurling, R. (2007). The moderator role of intuitive versus deliberative decision making for the predictive validity of implicit measures. *European Journal of Personality, 21*, 529–546.

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 105–120.

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

Rotteveel, M., & Phaf, R. H. (2004). Automatic affective evaluation does not automatically predispose for arm flexion and extension. *Emotion, 4,* 156–172.
Rudman, L. A. (2004). Sources of implicit attitudes. *Current Directions in Psychological Science, 13,* 79–82.
Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17,* 437–465.
Schmukle, S. C., & Egloff, B. (2006). Assessing anxiety with Extrinsic Simon Tasks. *Experimental Psychology, 53,* 149–160.
Schnabel, K., Banse, R., & Asendorpf, J. B. (2006). Employing automatic approach and avoidance tendencies for the assessment of implicit personality self-concept: The Implicit Association Procedure (IAP). *Experimental Psychology, 53,* 69–76.
Seibt, B., Neumann, R., Nussinson, R., & Strack, F. (2008). Movement direction or change in distance?: Self- and object-related approach–avoidance motions. *Journal of Experimental Social Psychology, 44,* 713–720.
Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. A., & Groom, C. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review, 115,* 314–335.
Solarz, A. K. (1960). Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. *Journal of Experimental Psychology, 59,* 239–245.
Sriram, N., & Greenwald, A. G. (2009). The Brief Implicit Association Test. *Experimental Psychology, 56,* 283–294.
Stahl, C., & Degner, J. (2007). Assessing automatic activation of valence: A multinomial model of EAST performance. *Experimental Psychology, 54,* 99–112.
Steffens, M. C., & Plewe, I. (2001). Items' cross-category associations as a confounding factor in the Implicit Association Test. *Zeitschrift für Experimentelle Psychologie, 48,* 123–134.
Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8,* 220–247.
Stroop, J. R. (1935). Studies on the interference in serial verbal reactions. *Journal of Experimental Psychology, 59,* 239–245.
Teachman, B. A. (2007). Evaluating implicit spider fear associations using the Go/No-Go Association Task. *Journal of Behavior Therapy and Experimental Psychiatry, 38,* 156–167.
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.
Teige, S., Schnabel, K., Banse, R., & Asendorpf, J. B. (2004). Assessment of multiple implicit self-concept dimensions using the Extrinsic Affective Simon Task. *European Journal of Personality, 18,* 495–520.
Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2008). Minimizing method-specific variance in the IAT: The Single Block IAT. *European Journal of Psychological Assessment, 24,* 237–245.
Teige-Mocigemba, S., Klauer, K. C., & Sherman, J. W. (2010). A practical guide to Implicit Association Tests and related tasks. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 117–139). New York: Guilford Press.



Wiers, R. W., Van Woerden, N., Smulders, F. T. Y., & De Jong, P. J. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology, 111,* 648–658.
Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107,* 101–126.
Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationships with questionnaire measures. *Journal of Personality and Social Psychology, 72,* 262–274.